DTU Library

Optimising Transport Decision Making using Customised Decision Models and Decision Conferences

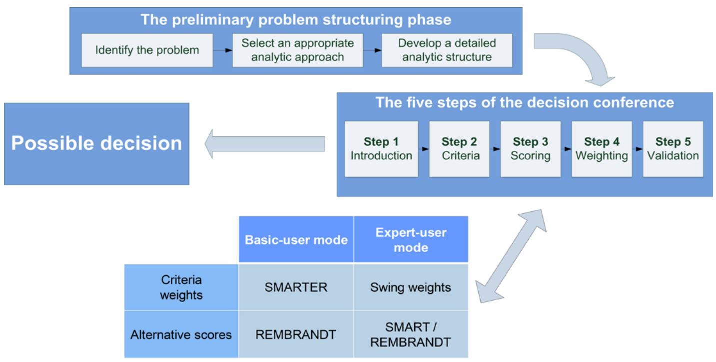
Barfod, Michael Bruhn

Publication date:
2012

Document Version
Publisher's PDF, also known as Version of record

Optimising Transport Decision Making using Customised Decision Models and Decision Conferences

PhD Thesis



Michael Bruhn Barfod
February 2012

DTU Transport
Department of Transport

Optimising Transport Decision Making using Customised Decision Models and Decision Conferences

PhD thesis

Michael Bruhn Barfod
Department of Transport
Technical University of Denmark

Supervisor:
Professor Steen Leleur
Department of Transport
Technical University of Denmark

PhD-2012-01
Department of Transport
Technical University of Denmark

Published by: DTU Transport
Bygningstorvet 116B
2800 Kgs. Lyngby
Denmark

Order at: http://www.transport.dtu.dk
Tel +45 45256500, Fax +45 45936412
transport@transport.dtu.dk

ISBN: 978-87-7327-228-2 (Electronic version)
ISBN: 978-87-7327-229-9 (Printed version)

Preface

The following thesis completes my Ph.D. study entitled "Optimising Transport Decision Making using Customised Decision Models and Decision Conferences". The study has been carried out at the Department of Transport (DTU Transport) at the Technical University of Denmark (DTU) in the period from January 2008 to December 2011 in a so-called ¾ time frame. The study has been partly financed by an institute scholarship, and has been supervised by Professor Steen Leleur.

During the study period I have been assigned various projects among others in collaboration with the Danish Road Directorate, the University of Lund and other partners in the Oresund Region. The experience that I have obtained from these projects has been invaluable and has contributed significantly to the case material presented in this thesis.

The process of writing a Ph.D. thesis has been an interesting, challenging and educating experience. The study has taken me to several places all over the world, where I have had the opportunity of sharing and discussing my ideas with other researchers and fellow Ph.D. students. Throughout the study I have followed courses on different universities both in Denmark and abroad as well as participated in several international conferences. At these occasions I have met many interesting and highly skilled researchers, who have provided me with constructive feedback on my work.

I would like to thank the people who have supported me throughout the last four years and contributed to the writing of this thesis. First of all I would like to thank my supervisor Professor Steen Leleur for encouraging me to apply for the Ph.D. scholarship and for his commitment and continuous support during my study. Special thanks also go to my colleagues, the present and former members of the Decision Modelling Group at DTU Transport: Kim Bang Salling, Anders Vestergaard Jensen, Sara Lise Jeppesen, and Inga Ambrasaite for helpful discussions, proof readings, co-authorships in some of the papers in this thesis, and good company on my travels to conferences and courses abroad.

Finally, my deepest thanks go to my family for their support and care over the years. Without the encouragement and help from my mother I would probably not have come this far today, and for this I am truly grateful. Last but not least I would like to thank my dear wife Premnet and my son Oliver for their love and care, and for keeping up with a sometimes forgetful and absent-minded husband and father.

Frederikssund, December 2011

Michael Bruhn Barfod

Abstract

The subject of this Ph.D. thesis entitled “Optimising Transport Decision Making using Customised Decision Models and Decision Conferences” is multi-criteria decision analysis (MCDA) and decision support in the context of transport infrastructure assessments. Despite the fact that large amounts of money are spent on examinations of transport infrastructure projects, such as traffic model calculations, environmental impact assessments (EIA), and public hearings, the results mainly express the outcome of the examinations in monetary terms in form of e.g. a benefit-cost rate (B/C-rate). This thesis is concerned with the insufficiency of conventional cost-benefit analysis (CBA), and proposes the use of MCDA as a supplementing tool in order to also capture impacts of a more strategic character in the appraisals and hence make more use of the often large efforts put in the preliminary examinations.

MCDA depends to a high degree on subjective preferences stated by the decision-makers as the methodology deals with impacts (or criteria) that are difficult to quantify or assign with a monetary value. As a result of this an examination process is proposed that can guide the decision-makers through the difficult task of assessing the impacts. Important for this process is that it should be based on appropriate methods and techniques, which are capable of modelling the decision-makers’ preferences and well as communicating the results.

The main focus of this Ph.D. study has been to develop a process and framework for providing valid, flexible and effective decision support in situations where complex decision problems concerning transport infrastructure projects are to be assessed. Throughout the study five papers have been produced laying the foundation with different case examples ranging from road and rail to bike transport projects. Two major concerns have been to propose an examination process that can be used in situations where complex decision problems need to be addressed by experts as well as non-experts in decision making, and to identify appropriate assessment techniques to be used in the decision process.

The first contribution of this Ph.D. study is a framework of MCDA techniques to be used in decision processes. Depending on which type of persons that is to be involved in the decision process different assessment techniques are proposed. Two main modes are in this respect relevant: a basic-user mode consisting of non-experts, and an expert-user mode consisting of professional and experienced users of the techniques. The second contribution of the study is an examination process that proposes how the appraisal of infrastructure

projects can be designed from the initial problem identification to the possible decision making. The process makes use of a preliminary problem structuring phase, and an intervention phase featuring the concept of a decision conference where decision-makers and multiple stakeholders have the possibility of interacting with the decision support model and thereby also influencing the results.

Based on the methodology and process developments throughout the thesis the following four main findings are presented:

1. The composite model for assessment (COSIMA) is an effective decision support system (DSS) for complex planning problems involving both monetary impacts and non-monetary criteria.

2. Direct rating using pair wise comparisons is found to be an appropriate MCDA approach for computing scores for alternatives while rank based approaches are appropriate for eliciting criteria weights from the decision-makers' preferences.

3. Decision analysis and decision conferences using MCDA are useful approaches for structuring and appraising large and complex decision problems with participation of relevant stakeholders and decision-makers.

4. The REMBRANDT technique with its better theoretical foundation can with a modified progression factor be recommended for practical use instead of the original AHP to derive decision-maker preferences.

In summing up, this Ph.D. thesis provides a broad foundation for further exploration and application of a MCDA based decision support framework. It is concluded based on the findings that MCDA ought to have a more widespread use in transport planning as several types of appraisal problems can be approached in an adequate way by making use of MCDA, where process and methodology is customised (optimised) in accordance with the actual case dealt with. A number of perspectives and future research possibilities are outlined related to both the applications of MCDA techniques and the decision process.

Abstract in Danish

Emnet for denne Ph.D. afhandling er multi-kriterie beslutningsanalyse (MCDA) og beslutningsstøtte i forbindelse med vurderinger af transportinfrastrukturprojekter. På trods af, at store pengesummer bruges på undersøgelser af transportinfrastrukturprojekter, såsom trafikmodelberegninger, miljøkonsekvensvurderinger (VVM), og offentlige høringer, udtrykkes resultaterne først og fremmest monetært i form af f.eks. en benefit-cost rate (B/C-rate). Denne afhandling beskæftiger sig med utilstrækkeligheden ved konventionel cost-benefit analyse (CBA), og foreslår anvendelse af MCDA som et supplerende redskab for også at medtage effekter af en mere strategisk karakter i vurderingerne og dermed gøre mere brug af de ofte store indsatser, som er gjort ved de indledende undersøgelser.

MCDA afhænger i høj grad af beslutningstagernes subjektive præferencer, da metodikken beskæftiger sig med effekter (eller kriterier), som er vanskelige at kvantificere eller tildele en monetær værdi. På baggrund af dette foreslås en vurderingsproces, som kan guide beslutningstagerne gennem den vanskelige opgave med at vurdere effekterne. Denne proces bør baseres på passende metoder og teknikker, som er i stand til at modellere beslutningstagernes præferencer samt formidle resultaterne på en tilfredsstillende måde.

Det primære fokus i denne Ph.D. afhandling har været at udvikle en proces og give nogle rammer for at udføre gyldig, fleksibel og effektiv beslutningsstøtte i situationer, hvor komplekse beslutningsproblemer vedrørende transportinfrastrukturprojekter skal vurderes. Igennem studieforløbet er produceret fem artikler omhandlende forskellige cases, som spænder fra vej og jernbane til cykeltransportprojekter. To problemstillinger har i denne sammenhæng været, at foreslå en undersøgelsesproces til brug i situationer, hvor komplekse beslutningsproblemer skal løses af eksperter såvel som ikke-eksperter, samt at identificere passende vurderingsteknikker til anvendelse i beslutningsprocesser.

Det første bidrag fra denne Ph.D. afhandling er et rammesystem af MCDA teknikker til brug i beslutningsprocesser. Afhængigt af hvilken type personer, der skal inddrages i beslutnings-processen foreslås forskellige teknikker til vurderingen. To hovedniveauer er i denne henseende relevante: et basis-bruger niveau til benyttelse af ikke-eksperter, og et ekspert-bruger niveau til benyttelse af eksperter og erfarne brugere af teknikkerne. Det andet bidrag fra afhandlingen er en undersøgelsesproces, der foreslår, hvordan vurderingen af infrastrukturprojekter kan designes fra den indledende problemidentifikation til den endelige beslutning tages. Processen gør brug af en indledende problemstruktureringsfase

samt en fase, hvor beslutningstagerne og eventuelle interessenter har mulighed for at interagere med den understøttende beslutningsmodel og dermed påvirke resultaterne.

Baseret på metode- og procesudviklingen igennem hele studieforløbet kan de følgende fire vigtigste resultater præsenteres:

1. Den sammensatte model til vurdering (COSIMA) er et effektivt beslutningsstøttesystem til at vurdere komplekse planlægningsproblemer, der involverer både monetære effekter og ikke-monetære kriterier.

2. Parvise sammenligningsteknikker er en passende MCDA metode til beregning af scores for alternativer, mens rang-baserede tilgange er egnede til bestemmelse af kriterievægte ud fra beslutningstagernes præferencer.

3. Beslutningsanalyse og beslutningskonferencer, der gør brug af MCDA, er nyttige metoder til strukturering og vurdering af store og komplekse beslutningsproblemer, som relevante interessenter og beslutningstagere kan deltage i.

4. Ved benyttelse af parvise sammenligninger kan REMBRANDT teknikken med en modificeret progressionsfaktor anbefales til praktisk brug i stedet for den oprindelige AHP.

Denne Ph.D. afhandling giver et bredt grundlag for yderligere undersøgelser og anvendelse af MCDA-baseret beslutningsstøtte. Det kan på baggrund af resultaterne konkluderes, at MCDA kunne have en mere udbredt anvendelse inden for transportplanlægning. Mange typer vurderingsproblemer kan gribes an på en passende måde ved brug af MCDA, hvor proces og metode tilpasses (optimeres) i overensstemmelse med det konkrete case. En række perspektiver og fremtidige forskningsmuligheder er skitseret relateret til både applikationer af MCDA teknikker og beslutningsprocesser.

Table of contents

Abbreviations

The abbreviations found below are used throughout this thesis. They are presented by their name when first encountered in each chapter but will afterwards be referred to by their abbreviation. In the list below the abbreviations are presented in alphabetic order. The list is thought as a help for the reader if the need for refreshing one or more abbreviations after its introduction should occur.

Abbreviation	Full name
AHP	Analytic Hierarchy Process
B/C-rate	Benefit/Cost rate
CBA	Cost-Benefit Analysis
CI	Consistency Index
CR	Consistency Ratio
COSIMA	COmpoSIte Model for Assessment
DMT	Danish Ministry of Transport
DSS	Decision Support System
EIA	Environmental Impact Assessment
IRR	Internal Rate of Return
MAUT	Multi-Attribute Utility Theory
MAVT	Multi-Attribute Value Theory
MCA	Multi-Criteria Analysis
MCDA	Multi-Criteria Decision Analysis
NPV	Net Present Value
REMBRANDT	Ratio Estimations in Magnitudes or deci-Bells to Rate Alternatives which are Non-DominaTed
ROD	Rank Order Distribution
RoH	Rule of a Half

RP	Revealed Preference
SMART	Simple Multi-Attribute Rating Technique
SMARTER	Simple Multi-Attribute Rating Technique Exploiting Ranks
SP	Stated Preference
TRR	Total Rate of Return
TV	Total Value
WTA	Willingness To Accept
WTP	Willingness To Pay

Overview of papers

Paper 1: *Composite decision support by combining cost-benefit and multi-criteria decision analysis*
Co-authors: Assistant Professor Kim Bang Salling (DTU Transport) and Professor Steen Leleur (DTU Transport)
Journal: *Decision Support Systems* 51 (2011), pp. 167-175, Elsevier (ISI-indexed)
Case: The new Hornsherred connection, Denmark
Keywords: Decision support systems, decision analysis, cost-benefit analysis, multi-criteria decision analysis

Paper 2: *Examination of decision support systems for composite CBA and MCDA assessments of transport infrastructure projects*
Co-authors: Ph.D. Student Anders Vestergaard Jensen (DTU Transport) and Professor Steen Leleur (DTU Transport)
Presented at: The 20th International Conference on Multiple Criteria Decision Making 2009 (MCDM2009)
Journal: Y. Shi et al. (Eds.). *New State of MCDM in the 21st Century – Selected papers of the 20th International Conference on Multiple Criteria Decision Making 2009*. Lecture Notes in Economics and Mathematical Systems 648 (2011), pp. 167-176, Springer
Case: The Ostlänken railway line, Sweden
Keywords: Decision support systems, multi-criteria decision analysis, COSIMA, REMBRANDT

Paper 3: *Customised DSS and decision conferences*
Co-author: Professor Steen Leleur (DTU Transport)
Presented at: The 13th Euro Working Group on Transportation Meeting 2009 (EWGT2009)
Journal: Proceedings of the 13th Euro Working Group on Transportation Meeting 2009, University of Padova, Italy
Case: The Ostlänken railway line, Sweden
Keywords: Decision conference, decision support systems, multi-criteria decision analysis, REMBRANDT, swing weights

Paper 4: *An MCDA approach for the selection of bike projects based on structuring and appraising activities*

Presented at: The 24th European Conference on Operational Research 2010 (EURO2010)
Journal: *European Journal of Operational Research* 218 (2012), pp. 810-818, Elsevier (ISI-indexed)
Case: Prioritising bike projects (CPP), Denmark
Keywords: Decision analysis, problem structuring, multi-criteria decision analysis, REMBRANDT, decision support systems

Paper 5: *Scaling transformations in the REMBRANDT technique: a sensitivity examination of the progression factors*
Co-author: Professor Steen Leleur (DTU Transport)
Presented at: The 21st International Conference on Multiple Criteria Decision Making 2011 (MCDM2011)
Journal: *International Journal of Information Technology & Decision Making* (under review), World Scientific (ISI-indexed)
Case: The fixed link between Elsinore and Helsingborg, Denmark and Sweden
Keywords: Multi-criteria decision analysis, decision support systems, REMBRANDT

1 Introduction

Traditionally appraisals of transport infrastructure projects have mainly been limited to conventional cost-benefit analysis (CBA). This is especially true in Denmark where the Manual for Socio-economic Appraisal (DMT, 2003) sets the standard for the evaluations, but also in many other European countries. The recent years' political tendencies aiming at greater considerations with regard to both the environment and other "soft" impacts have, however, indicated that the CBA methodology no longer is sufficient for the appraisals. Hence, a need has arisen for the inclusion of other impacts besides the economic in the overall appraisals. At present no guidelines, however, exist for how these "soft" impacts can be combined with the conventional analysis into a more comprehensive type of appraisal. In this respect multi-criteria decision analysis (MCDA) is proposed as an additional methodology.

Assessments using MCDA will to a high degree depend on the decision-makers' preferences as the methodology often deals with criteria that are very difficult to quantify or assign with a monetary value. Thus the criteria need to be assessed using subjective judgments. As a result of this there is a need for the development of a methodology that can guide the decision-makers through the process of assessing the impacts. Important for this process is that it is carried through using appropriate methods which are capable of modelling the decision-makers' preferences as well as communicating the results to third parties in order to ease the acceptance of the final decision. This sets some high demands to the methodology used as it will be necessary to adjust the combination of methods specifically to each single decision problem.

1.1 Research questions

An efficient and optimal infrastructure is of great importance to the society. In order to obtain this it is important to make comprehensive and long-term appraisals of those infrastructure projects that are planned for. For this reason the decision process should be capable of handling these transport related problems which are often very complex and politically loaded issues. The main scope of this thesis has been to examine whether it is possible on the basis of a theoretical and practical approach to make use of processes and methods both from the field of transportation but also from other scientific fields, sectors and countries which also work with the refinement of decision processes. Thus, it is examined whether aspects and perceptions exist that should be included in the appraisal methodology in order to optimise it to provide more comprehensive and transparent results in preparation for use in actual decision making.

The focus of this thesis is framed by the following two main research questions:

I. Is it possible to propose an examination process that can be used in situations where complex decision problems need to be addressed by technicians as well as decision-makers and citizens?

II. How can the methodologies and techniques made use of within the examination process be optimised to meet the specific decision task in hand?

The concept of optimising the examination process should in this context be perceived as making the process as effective as possible by finding the best compromise between the methodologies and techniques available. The questions are addressed in details in the five papers which form this thesis. More specifically each of the papers addresses a specific sub-question of the main research questions. These are:

- *Paper 1:* Can comprehensive appraisals taking into account both monetary impacts and non-monetary criteria of a decision problem be operationalised to a decision support system that can inform the users in terms of both interaction and interpretation of the results?

- *Paper 2:* Which of the two decision support systems, COSIMA and REMBRANDT, is the most appropriate for conducting composite appraisal of transport infrastructure projects in terms of the level of information to the decision-makers?

- *Paper 3:* Can a structured decision making process be outlined which can take all important aspects into account and at the same time be transparent both to the participants and the public?

- *Paper 4:* Can the theory of decision analysis be useful to structure a complex decision problem concerning transport infrastructure issues, and can a set of guidelines be formulated for the appraisal of the decision problem using multi-criteria decision analysis?

- *Paper 5:* What is the influence of the progression factors for the scaling transformation in the REMBRANDT technique towards the final outcome of a decision analysis, and can a revision of these factors be proposed in order to make the results of the technique more acceptable?

1.2 Structure of the thesis

The thesis is structured as follows:

Chapter 1 presented the purpose and outline of this Ph.D. thesis. This included an outline of the main research questions, followed by an outline of the content of the thesis.

Chapter 2 presents an overview of the main methodologies addressed by this thesis, namely cost-benefit analysis (CBA) and multi-criteria decision analysis (MCDA). The methodologies are elaborated to an appropriate extent for the thesis and their main strengths and weaknesses are discussed.

Chapter 3 presents the modelling approach made use of in the thesis. The modelling approach takes its basis in the value measurement theory of MCDA. Thus this theory will firstly be outlined after which the different methods applied are presented, namely the SMART technique (simple multi-attribute rating technique), the AHP technique (analytic hierarchy process) and the REMBRANDT technique (ratio estimations in magnitudes or deci-bells to rate alternatives which are non-dominated). Next, the composite model for assessment (COSIMA) is outlined presenting an approach for combining CBA and MCDA. Finally, the concept of decision conferences is outlined.

Chapter 4 presents the purpose and findings of each of the five papers produced during the Ph.D. study. The cases are described and the main findings of each paper are presented. The full papers are enclosed in the end of the thesis.

Chapter 5 presents the findings relating to assessment techniques and examination process that can be made on the basis of the work made in the five papers, and a discussion of the validity of the findings is made.

Finally, **Chapter 6** presents the conclusions of this Ph.D. thesis which lead to answering the research questions outlined in the present chapter. Moreover, perspectives on future work within the research area are given.

2 Socio-economic evaluation

In order to conduct a socio-economic evaluation of a project or initiative it is necessary to apply an evaluation methodology that is suitable for an appropriate handling of the issue. This chapter presents the two appraisal methodologies of cost-benefit analysis (CBA) and multi-criteria decision analysis (MCDA) respectively and accounts for their differences with respect to the underlying theory.

2.1 Cost-benefit analysis

Cost-benefit analysis (CBA) is a widely applied method for evaluating the “goodness” of public investments as well as for ranking alternative investments. The basic feature of the methodology is the comparison of costs and benefits, which are all measured on the same monetary scale (Dasgupta and Pearce, 1978). The main purpose of the analysis is to improve decision making – to enable those responsible for decisions to choose projects with higher net benefits over those with lower net benefits.

The CBA of a public investment can to some extent be compared with an economic analysis carried out by a private company. Such a private company will conduct careful analyses and make decisions that maximise its future revenue. In such analyses, the private company will e.g. use the product’s sales price as a measure of the benefit and the price of production as a measure for the costs (Gissel, 1999).

2.1.1 Basic principles of CBA

The theoretical basis of CBA rests upon the micro-economic concept of welfare theory. The fundamental assumption in micro-economic theory is that of a *rational consumer* (Gissel, 1999). This means that given the choice set available to the consumer, he will make choices in a way that maximises his own welfare (or utility) which is generally assumed to be represented by a utility function, u (Dasgupta and Pearce, 1978).

As society to some extent consists of its individuals, it seems natural to observe the social change in welfare from a given investment as an aggregate value of the individual utility gains and losses (Gissel, 1999). Hence, when constructing a social welfare function, the natural approach is to aggregate the individual utility functions into one function representing the welfare to society as a whole. This welfare function has much the same interpretation as the individual utility function: the alternative, which maximises the social welfare function, is preferred. One choice of welfare function is to express the social welfare, W, as an un-weighted sum of the individual utility gains and losses:

$$W = u_1 + u_2 + \cdots + u_n \tag{2.1}$$

where n is the number of individuals affected.

Such a welfare function is said to be additively separable with respect to individual utilities, and it is called a utilitarian welfare function (Gissel, 1999). This type of welfare function is the traditional approach in CBA and basically means that the gain of one person can compensate for the loss of another person, and that equal weights are assigned to all groups in society. Or, in other words, a benefit of 100 DKK has the same weight with respect to society's welfare whoever receives it or loses it – rich or poor. Hence, questions on equity or distributional aspects need to be considered separately.

As society's welfare is based on individual utilities, such values should be derived on the basis of individual preferences where possible. Accordingly, the value of a benefit should be derived as the amount of money an individual is willing to give up to obtain the benefit, and, similarly, the value of a cost element should be derived as the amount of money an individual is willing to accept as a compensation.

Such willingness-to-pay (WTP) and willingness-to-accept (WTA) values are generally not derived by asking individuals directly – such questions would be too difficult to answer and, in some cases, could involve the risk of strategic answers (e.g. if the individual suspects that the WTP for the travel time reduction would not be completely hypothetical and that he will be charged extra according to the WTP he states). Instead a variety of indirect methods can be used.

One such method is the stated preference (SP) approach, where individuals through their answers to hypothetical choice situations indirectly express their WTP or WTA. This is a widely applied method for assessing values of non-marketed effects.

The revealed preference (RP) technique involves observing and examining the actual behaviour of people in situations where they have a choice between, for example, an expensive and fast travel possibility and a less expensive but more time consuming possibility. The difference between RP and SP is that the SP concerns hypothetical behaviour as people are asked their opinion about various hypothetic choice situations. The SP type of examination has other problems associated with it than is the case with the RP examination. Where influential factors can be difficult to neutralise in RP, these can neatly be defined and neutralised with a well designed questionnaire used in the SP. On the other hand, people may respond more defiantly to hypothetic situations than to real situations (Næss, 2006; Leleur, 2000).

The SP and RP methods are not always applicable and other valuation methods exist. These can be categorised into two groups: valuation methods based on individual preferences and methods based on costs. For a more thorough description see e.g. Gissel (1999) or Leleur (2000). Valuation methods based on costs do not necessarily reflect the actual WTP or WTA, and, in many situations, they are merely estimates on a minimum WTP. Therefore SP and RP should be preferred if (practically and financially) possible.

2.1.2 Rule-of-a-half

The WTP measure can be illustrated graphically by considering a demand curve. A demand curve expresses society's demand (i.e. its WTP) for a commodity as a function of its price. As an example consider the quantity of travellers, *Q*, for travel on a given road section. This quantity depends among others on the travel time and comfort. Assume that these factors may be combined into a total price, *P*.

The demand curve for trips on the road section is illustrated in Figure 2.1. In the initial situation there are *Q* travellers who all experience a price, *P*. Assume now that an infrastructure repair enables the travellers to increase their speed hence reducing the travel time – and thereby the price. If the price reduces to *P'* the demand will increase to *Q'*. The existing *Q* travellers will experience a cost reduction of *P-P'* implying that each of the existing travellers will acquire a benefit of this size. Hence, the total benefit to existing travellers may be described by area *A* in Figure 2.1.

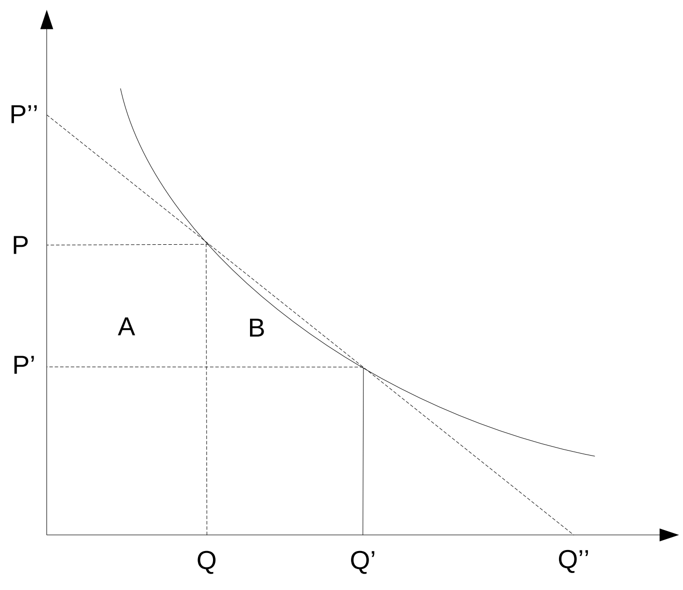


Figure 2.1. Demand curve – the sum of the areas A and B describes the total user benefit obtained by reducing the price, P

Consider the *Q*+1'th traveller. Before the reduction in travel time he did not travel because he was only willing to "pay" (or endure) a price slightly less than *P*. After the cost reduction he experiences a cost of *P'*, and as his WTP is higher, he experiences a benefit equal to the vertical distance between *P* and the demand curve (describing the WTP). Similar reasoning may be applied to traveller *Q*+2 up to traveller *Q'*, implying that the benefit accruing to the new travellers may be found as the area *B* in Figure 2.1.

The total benefit of the price change may thus be described as the area under the demand curve between the "before" and the "after" prices (*P* and *P'*) – i.e. the sum of the areas *A* and *B*. This can be expressed by what is normally referred to as the 'Rule-of-a-Half' (RoH).

$$RoH = (P - P') \cdot Q + \frac{1}{2} \cdot (P - P') \cdot (Q' - Q) = \frac{1}{2} \cdot (P - P') \cdot (Q + Q') \qquad (2.2)$$

The expression implicitly assumes that there is a linear relationship between the price and demand. If this is not the case, and the demand curve is convex to the origin, then the RoH will tend to overstate the benefits. With very small changes in price, the inaccuracy is, however, not significant.

2.1.3 Investment criteria

Transport infrastructure projects are characterised by having impacts that will change over the years of the evaluation period. Generally, a construction phase, which has costs in the opening year, will be replaced with benefits in the following years, due to a continuously increasing traffic that will grow steadily (Leleur, 2000). This development is depicted in Figure 2.2 (note that the figure depicts non-discounted benefits and costs).

Different types of projects of varying sizes are normally characterised by differences in their development of future benefits. However, in order to aggregate the streams of costs and benefits into a single value, which reflects the profitability of the project, existing economic index values can be applied. Such index values are useful for socio-economic analysis, but they should be selected in accordance with their valid applicability and applied based on the availability of data in the appraisal task in hand (Ibid.).

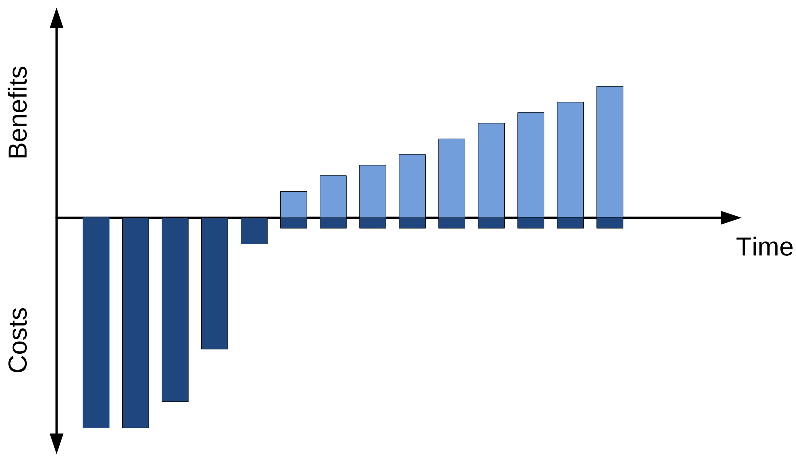


Figure 2.2. Development of costs and benefits over the years (adapted from Leleur (2000), p. 100)

The following which makes use of the work performed by Dasgupta and Pearce (1978), Gissel (1999), Banister and Berechman (2000) and Leleur (2000) includes a brief description of the net present value, the internal rate of return, and the benefit-cost rate, which are the three most widely used investment criteria in traffic planning.

Net present value

The net present value (NPV) criterion requires that (2.3) is to be evaluated for all investment alternatives. For every investment, the streams of benefits and cost are aggregated into one single value, which indicates the profitability of the project or initiative. A minimum demand for this is that NPV is positive:

$$NPV = \sum_{t=0}^{T} \frac{B_t - C_t}{(1+r)^t} \tag{2.3}$$

T is the total number of years in the evaluation period of the project or initiative, B_t is the amount of benefits in year t, C_t is the amount of costs in year t and r is the discount rate.

The principal content of the NPV calculation consist of the different time-dependent weights attached to the time-displaced benefits and costs using the discount factor $(1+r)^{-t}$, where a fixed discount rate is normally applied with $r > 0$. The higher values of r and t, the lesser added contribution from the discounted value, see Figure 2.3.

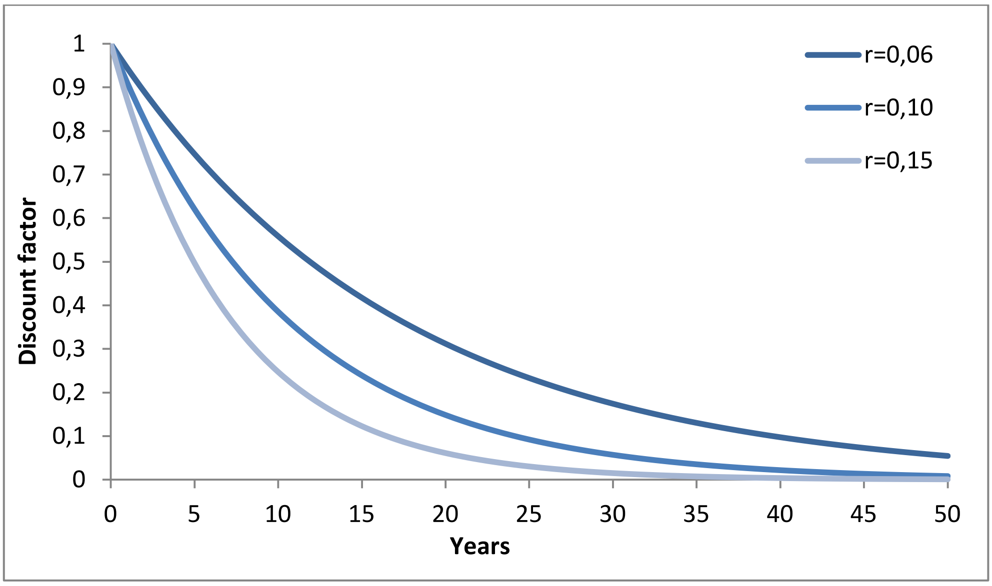


Figure 2.3. The discount factor $(1+r)^{-t}$ as a function of t and the discount rate r (adapted from Leleur (2000), p. 101)

The actual value of the discount rate is an expression of the emphasis on benefits in the near future compared with benefits in a more distant future. Due to the types of projects associated with the benefit types, a low calculation rate will favour larger projects with a long project life, while a high rate will lead to a comparatively higher profitability of projects lesser in costs and size. In Denmark the discount rate has varied from 7 % in year 2000 to 6 % in year 2003 and to 5 % in 2009, and the rate varies across Europe.

When conducting a NPV calculation, a base year must be determined for price level reference. No attention is in this respect paid to inflation, but account can be taken of forecasted growth in real terms of some of the benefit components unit prices. Using the NPV as the decision criterion implies that all projects with a positive NPV should be carried through. However, if there are only limited financial resources and not all projects from a project pool with a positive NPV can be implemented, the relative value of these projects needs to be considered in order to rank them.

Internal rate of return

The purpose of the internal rate of return (IRR) is to determine the rate, *i*, which balances the cost and benefits streams, i.e. equates the present value of the stream of expected benefits in excess of cost to zero:

$$\sum_{t=0}^{T} \frac{B_t - C_t}{(1+i)^t} = 0 \qquad (2.4)$$

The higher the rate *i*, the better the examined project or initiative. An advantage of this investment criterion is that a calculation rate is not needed as when calculating NPV. An uncertainty with the IRR method is, however, that it is the solution to a polynomial equation with several roots, which cannot always easily be sorted out.

While the IRR method is effective in deciding whether or not a project is worth undertaking, it is difficult to utilise in ranking projects and in deciding between competing alternatives if budget constraints are an issue. It is not unusual for rankings established by the IRR method to be inconsistent with those of the NPV and benefit/cost rate investment criteria.

Benefit/cost rate

The benefit/cost rate (B/C-rate) criterion is defined as the present value of benefits divided by the present value of costs, and is given by:

$$B/_C = \frac{\sum_{t=0}^{T} \frac{B_t}{(1+r)^t}}{\sum_{t=0}^{T} \frac{C_t}{(1+r)^t}} \qquad (2.5)$$

A minimum demand for the criterion is that B/C > 1. When applying the B/C-rate, it must be noted that the criterion calculates the discounted benefits per discounted investment unit. Thus, if a comparison is carried out among a group of projects that differ in size and investment demand, the B/C-rate will not determine the project with the numerically largest net benefits as is the case with the NPV. On this basis, however, it is relevant to apply the B/C-rate in connection with priority studies under a budget.

2.1.4 Strengths and weaknesses of CBA

The strengths of the CBA methodology are quite convincing and well-known due to its many applications. They are for this reason only briefly mentioned in this section and may be categorised as follows:

- Transparency
- Comparability / consistency
- Ignorance revelation (through systematic collection of information)

First, the CBA is transparent in the sense that it converts all impacts into an absolute monetary measure as it is desirable to be able to sum up all aspects of the decision problem into a single value describing the social profitability. Second, the CBA provides a methodological tool for comparing projects and/or alternatives which ensures that the same result is reached no matter who conducts the analysis. This makes it a powerful decision support tool in the planning process as the values of cost and benefit elements are consistent between investments and over time. Third, the CBA requires the collection of detailed information of financial as well as social costs and benefits. This gathering of information improves the basis on which the decision is made and may give valuable insight into the level of ignorance regarding important aspects of the evaluated project or policy (Leleur et al., 2004).

Naturally there are also different issues associated with the CBA method. These can be categorised as follows:

- “False” transparency
- Practical measurement problems
- Inter-generational equity
- Social equity

It is difficult to maintain consistency between the theoretical assumptions of the CBA method and the practical application of it. This is mainly due to the fact that there may be problems involved when estimating unit prices for non-marketed impacts such as travel time savings, emissions, accidents and noise. In practice, therefore, compromises are often made on the valuation of such non-marketed impacts, implying that the resulting unit prices are generally of a subjective nature – without such subjectivities being visible in the evaluation. This is a problem with the CBA method since the presentation of a single evaluation measure thus implies a “false” sense of objectivity.

What is seen by most economists as one of the great advantages of CBA, namely its great transparency, is argued by others as the exact opposite: All financial, environmental and social considerations are reduced to a single point estimate – thereby shielding the results behind a technical mystique. This disagreement can be argued as being a matter of taste, but it is a real problem if the general public perceives the evaluation method as some kind of “black box” (Gissel, 1999). Moreover, considering transport infrastructure appraisals, Salling

(2008) argues that basing the final decision making on a single point estimate such as the B/C-rate is problematic due to the uncertainties behind the traffic forecasts and cost estimates.

A study by Beukers et al. (2012) has by use of interview sessions with different focus groups – namely plan makers, CBA advisors, CBA makers, CBA testers, funding applicants, lobbyists and academics – examined how the CBA as a tool is perceived. The study concludes that CBA and its process are perceived as rather problematic and characterised as frustrating when applied to assess complex infrastructure plans among others due to reasons of deficient communication, fear of not having included all relevant effects, and time pressure in the process. In addition to this Osland and Strand (2010) notes that politicians seem to give relative little weight on the results of CBA in the decision making process.

Moreover, there are impacts which can hardly be quantified or for which it is difficult or even impossible to estimate unit prices. These are especially impacts of a more long-term and/or strategic nature – as for example many environmental impacts (Engelbrecht, 2009).

An important philosophical and moral problem in the evaluation of long term impacts is that of the present generation valuing an impact, which they may not live to experience. This means that they are valuing such impacts on behalf of the future generations, and costs and benefits that are more than thirty years away become almost without value when discounting at normal rates. Hence, long-term costs, such as e.g. environmental resource depletion may be effectively ignored in a CBA (Næss, 2006; Goklany, 2009). Discounting therefore discriminates against future generations by saying that future costs are worth less than present costs, and that present benefits are worth more than future benefits.

As noted by Ackerman and Heinzerling (2002) the logic behind discounting derives from the logic of money – that a person would prefer to receive money now rather than the same amount in the future (the time preference rate is positive). This is according to Beder (2000) because:

1. Money obtained now can be invested and earn interest
2. People tend to be impatient (they want to enjoy benefits sooner and costs later)
3. The person might die before he or she gets the money
4. One cannot be sure of getting the money in the future
5. People in the future will probably be better off; money will not be worth as much then

Seen from the society's point of view, it is more the number and types of individuals receiving a given benefit, which matters, and not whether it is a specific person. Hence, the idea that someone would like to consume now rather than in the future is not applicable to public goods, which can be enjoyed now and in the future. Also, the risk of one person dying before he or she gets the benefit is of no relevance if this person is just "exchanged" by another (as will be the case for a number of costs or benefit elements accruing over time). Any positive discount rate devalues future costs or benefits and this disadvantages future generations with respect to today's decisions. The logic of money – and in this respect the logic of discounting – may thus seem inappropriate when evaluating certain types of costs and benefits. This is especially the case for long term environmental impacts (Næss, 2006; Goklany, 2009).

The final problem with CBA concerning social equity can be divided into the three following issues relating to: individual welfare measurements, aggregating these individual measurements into one of social welfare, and the fact that no actual payment takes place.

Methods based on individuals' WTP are often used when valuing costs and benefits. People's WTP, whether measured directly or inferred in some way, is intimately linked with their ability to pay. Therefore the market can be seen as a system which advantages those most able to pay. The problem then arises that what people state their WTP for in such cases is a good conscience, not their own valuation of the good itself (Kahneman and Knetsch, 1992). Hence, using the market, whether an actual market or a contrived one, tends to produce values that reflect the existing distribution of income (Ackerman and Heinzerling, 2002).

In its conventional form CBA is about aggregated (and un-weighted) costs and benefits and does not deal with the issue of how they are distributed – although this is of prime concern when considering equity. As long as the sum of benefits outweighs the sum of costs (no matter who or how few people get the benefits and who or how many people suffer the costs) the society as a whole is assumed to be better off. It can be argued that in principle the CBA does not presuppose that individuals are treated anonymously – that is with equal weight in the aggregation of individual welfare into a measure of social welfare. In theory, one could aggregate individual welfare measures in a way (i.e. with weights) reflecting relevant equity concerns. However, anonymous aggregation has become the default in CBA no unique set of "equity weights" exists (Gissel, 1999).

Although the CBA method rests on the aggregation of individuals' WTP, no actual payment takes place and no actual redistribution of money is the result of this (Alcock and Powel, 2011). Hence, based on equity consideration the socio-economic results from the CBA could be argued as being rather hypothetical.

The conclusions on the discussions on the different strengths and weaknesses of CBA must be that the method is a reasonable appraisal methodology for projects where relevant cost and benefit elements can be monetised. However, if important impacts of the project cannot be given a direct monetary value – or if it is too costly to derive estimates of such values – the CBA of the remaining costs and benefits gives no real information on the total social value of the project.

2.2 Multi-criteria decision analysis

The field of multi-criteria decision analysis (MCDA) is concerned with the design of mathematical and computational tools to support the subjective evaluation of a finite number of decision alternatives under a finite number of performance criteria, by a single decision-maker or by a group of decision-makers and/or stakeholders.

A decision is in this respect defined as a choice out of a number of alternatives, and the choice is made in such a way that the preferred alternative is the “best” among the possible candidates. However, the decision process which precedes the choice is not always easy. Usually there are several ways to assess the alternatives and there is no alternative which outranks all the others under each of the performance criteria. Thus, the decision-maker does not only have the task to assess the performance of the alternatives in question under each criterion, he also has to weigh the relative performance of the criteria in order to arrive at a global assessment. Moreover, in a group of decision-makers each participant faces the question of how to assess the quality of the other participants and their relative power positions before an acceptable compromise solution emerges.

Many decisions take a long period of preparations, not only in a state bureaucracy concerning a decision about a large traffic investment, but sometimes also in a small organisational unit like a family. As soon as a problem has been identified which is sufficiently mature for action, a decision-maker is appointed or a decision making body is established. The choice of the decision-maker or the composition of the decision making body usually emerges as the result of a series of negotiations and reflects the strength or the influence of various parties. In general, the participants are also selected on the basis of their ability to assess at least some of the possible alternatives under at least some of the criteria. In the work of the decision making group the relevant criteria may have been prescribed and the relative importance of the criteria may have been formulated in vague verbal terms. Hence, the assessment of the alternatives under the prescribed criteria is left to the experts in the group, but the weighting of the criteria themselves is felt to be the prerogative or the responsibility of the authorities who established the group. During the deliberations it may happen that new alternatives and/or new criteria emerge and that the composition of the decision making body changes because new expertise is required. Nevertheless, there may be a clear endpoint of the decision process, in a particular session of the decision making group where each of the participants express their assessment. At this moment MCDA plays a significant role.

2.2.1 Objectives of MCDA

Methods for MCDA have been designed in order to select a preferred alternative, to classify the alternatives in a small number of categories, and/or to rank the alternatives in a subjective order of preference. Scanning the literature (von Winterfeldt and Edwards, 1986;

Stewart, 1992; Keeney and Raiffa, 1993; Lootsma, 1999; Saaty, 2001; Belton and Stewart, 2002; Edwards et al., 2007; Goodwin and Wright, 2009) it is found that MCDA usually has some or all of the following objectives:

1. **Improvement of the satisfaction with the decision process.** MCDA urges the decision-makers to frame the decision problem and to formulate the context explicitly. Next, MCDA structures the problem because the decision-makers are requested to list the alternatives and the criteria, and to record the performance of the alternatives under each of the relevant criteria, either in their original physical or monetary units or in verbal terms. Moreover, MCDA aids the decision-makers in the formulation of the criteria because it shows the priorities and values which may be deeply hidden in the back of their mind. In fact, it makes the criteria operational. MCDA also supports the decision-makers in the evaluation of the alternatives because it shows the subjective values of the performance of the alternatives within the context of the decision problem. Finally, MCDA eliminates or reveals the hidden objectives of certain participants in a group (the hidden agenda), and it reduces the effects of certain discussion techniques. It reduces the dominant role of participants with strong verbal skills, for instance, so that the silent majority has a proper chance to weigh the pros and cons of the alternatives and to insert their judgment in the decision process. In short, MCDA enhances the communication in the group.

2. **Improvement of the quality of the decision itself.** MCDA enables the decision-makers to break down a decision problem into manageable portions and to express a detailed judgment. The decision-makers are not easily swept away by the performance of some alternatives under one or two criteria only, but they keep an eye on the performance of all alternatives under all criteria simultaneously. Moreover, MCDA may propose a compromise solution in a group of decision-makers, possibly after several rounds of discussion, with due regard to the relative power positions of the group participants. The general experience is that MCDA may come up with an attractive proposal in an early stage of the decision process. The decision-makers, however, need several rounds of discussion before they accept the proposed alternative.

3. **Increased productivity of the decision-makers** – more decisions per unit of time. This objective ranks high on the agenda of decision-makers who are repeatedly involved in processes to evaluate the performance of certain types of projects. The criteria and alternatives are rather similar, from one session to the next, so that MCDA could be used to save time and energy.

MCDA also has some drawbacks. It introduces a formalised style of working, possibly in cooperation with an analyst and a facilitator using a computer or even a network of computers. This constitutes an extra burden for a decision-maker and particular for a group of decision-makers.

2.2.2 Research in MCDA

The research in the field of MCDA follows a number of distinct approaches in order to support the decision-makers in their attempts to identify a preferred alternative, to classify the alternatives, and/or to rank them. Based on von Winterfeldt and Edwards (1986), Bell et al. (1988) and Roy and Vanderpooten (1996) the approaches can be divided in four main categories:

1. **The descriptive approach** simply tells how decision-makers actually behave when they are confronted with the choice between several alternatives under conflicting viewpoints and how MCDA contributes to the decision process. Many studies are concerned with individual and collective decision making, with an analysis of the rationality of decision-makers in various cultural contexts, and with the identification of hidden objectives. Many specialists in this research area hesitate to believe that MCDA is compatible with the style of decision making of human beings. The rigidity and the formality of MCDA are mostly in conflict with the heuristic tactics applied by human beings during the decision process.

2. **The normative approach** tells how decision-makers should behave and how MCDA should work, via logical rules which are based upon certain fundamental axioms such as the transitivity of preferences. The typical products of the approach are Multi-Attribute Value Theory (MAVT) and Multi-Attribute Utility Theory (MAUT), to be used for decision problems with certain and uncertain outcomes respectively. The key element of MAVT is the concept of a value function which represents the degree to which the alternatives satisfy the objective of the criterion under consideration. In MAUT the user is requested to construct a utility function by choosing the monetary equivalent for a lottery or the equivalent lottery for a given amount of money. Despite the firmness of their axiomatic foundations these methods are not easy to use. The verbal probability estimates expressed by human beings, for instance, seem to be notorious poor. Furthermore, the rational axioms of preference judgment are not always obeyed by human beings. Cyclic judgment, for instance, can easily occur under qualitative criteria, when human beings carry out a dynamic search for a proper perspective on the alternatives.

3. **The prescriptive approach** tells how decision-makers could improve the decision process and the decisions themselves, and how MCDA could support such a process. Key elements in the approach are the modelling of human judgement, the identification of preference intensities, the aggregation processes, and the design of decision support systems. The approach can typically be regarded as a branch of Operations Research because it is also concerned with ad hoc models and algorithmic operations in order to support actual decision making. On the one hand, it may be successful when it concentrates on the needs of the users and on the potential benefits of mathematical analysis and information processing. On the other hand, it is always in danger when it ignores the unpredictability and the hidden agenda of the decision-makers.

4. **The constructive approach** questions the existence of a coherent, well-ordered system of preferences and values in the decision-makers' mind, as well as the idea that MCDA should correctly comprehend such a pre-existing system. These are the more or less tacit assumptions underlying the normative and prescriptive approach. The constructivists, however, advocate that the decision-maker and an analyst should jointly construct a model of the system, at least as far as it is relevant in the actual situation. In other words, the decision-makers' preferences and values, initially unstable and unpredictable or even non-existent, will be shaped by MCDA. This implies that the results of the analysis may be highly dependent on the analyst and on the method employed.

This short summary is a simplification of the basic philosophies behind the respective approaches. For a more thorough exposition of the descriptive, the normative and the prescriptive approach see von Winterfeldt and Edwards (1986) and Bell et al. (1988). The constructive approach is followed by the French school in MCDA (usually referred to as multi-criteria analysis – MCA), and it has been extensively described in the works of Bernard Roy (see e.g. Roy and Vanderpooten (1996) for an overview). The normative approach is usually referred to as the American School in MCDA. It is, however, unclear in the literature whether the methods which have been designed in the French and the American school respectively satisfy the typical needs of the French and American decision-makers. Research in MCDA is usually concerned with various methods within the framework of the respective approaches, not with the differences between styles of management and decision making in various parts of the world.

This thesis is only concerned with a restricted number of methods for MCDA. Because human beings normally express their preferences in terms which reveal gradations of intensity (indifference, weak, definite, strong or very strong preference), the thesis is limited to cardinal methods on the assumption that the preference information which is obtained in the respective elicitation processes constitutes ratio or difference information. Occasional the thesis turns to ordinal methods where the decision-makers merely rank-order their preferences. This thesis henceforth concentrates on two main methods and their applications:

1. **The Simple Multi-Attribute Rating Technique (SMART)**. The performance of the alternatives under the respective criteria, evaluated via a direct rating process, is expressed in grades on a numerical scale.

2. **The Analytic Hierarchy Process (AHP)**. The alternatives are considered in pairs. Their relative performance can equivalently be expressed as a ratio of subjective values (additive AHP) or as a difference of grades (multiplicative AHP – the REMBRANDT technique)

These two main methods form the basis for the further work in this thesis. The methods will be elaborated in Chapter 3.

3 Modelling approach

The modelling approach presented in this thesis takes its basis in the value measurement theory of multi-criteria decision analysis (MCDA). Thus the theory of value measurement, which is the basis for the later applied MCDA techniques, will first be outlined. After this the techniques of SMART, AHP, and REMBRANDT are presented in the sequence of their development. Finally, the COSIMA technique, which deals with the issue of combing cost-benefit analysis (CBA) and MCDA into one single measure of attractiveness, is presented, and the chapter concludes with a proposal for an approach to ease the decision making process.

3.1 Value measurement

The purpose of value measurement theory is to produce a means of associating a real number with each alternative in an assessment, in order to construct a preference order of the alternatives consistent with decision-maker value judgments. In other words, it is desirable to associate a number or value, *V(a)*, with alternative, *a*, in such a way that *a* is assessed to be preferred to *b*, taking all criteria into account, if and only if *V(a)* > *V(b)*. This also implies indifference between *a* and *b* if and only if *V(a)* = *V(b)*. Note, that the preference order implied by any such value function must constitute a complete weak order or pre-order (Belton and Stewart, 2002), i.e.:

> ***Preferences are complete****:* For any pair of alternatives, either one is strictly preferred to the other or there is indifference between them.
>
> ***Preferences and indifferences are transitive****:* For any three alternatives, e.g. *a*, *b* and *c*, if *a* is preferred to *b*, and *b* is preferred to *c*, then *a* is preferred to *c*, and similarly for indifference.

The value measurement approach thus constructs preferences which, in the first instance are required to be consistent with a relative strong set of axioms. However, it is important to note, that in practice value measurement will not be applied with such a literal and rigid view of these assumptions. The construction of a particular value function does impose the discipline of coherence with these “rationality assumptions”, but the results of and the conclusions from the value function will be subjected to intensive sensitivity analyses. The end result will generally be much less rigidly precise than may be proposed by the axioms.

Within the value measurement approach, the first component of preference modelling (measuring the relative importance of achieving different performance levels for each identified criterion) is achieved by constructing "marginal" (or "partial") value functions, $v_i(a)$, for each criterion. A fundamental property of the partial value function is that alternative *a* is preferred to alternative *b* in terms of criterion *i* if and only if $v_i(a) > v_i(b)$. Similarly, indifference between *a* and *b* in terms of this criterion exist if and only if $v_i(a) = v_i(b)$. Thus the partial value function satisfies the definition of a preference function; see Belton and Stewart (2002). However, the partial value functions will in addition need to model strength of preference in some sense, so that stronger properties than simple preservation of preference ordering will in general be needed.

Value function methods produce the assessments of the performance of alternatives against individual criteria, together with inter-criteria information reflecting the relative importance of the different criteria, w_i, to give an overall evaluation of each alternative indicative of the decision-makers preferences. The simplest and most widely used form of value function method is the additive model (von Winterfeldt and Edwards, 1986):

$$V(a) = \sum_{i=1}^{m} w_i v_i(a) \tag{3.1}$$

Considerably more complicated in appearance, but as easy to use, is the multiplicative model (Ibid.):

$$V(a) = \prod_{i=1}^{m} [v_i(a)]^{w_i} \tag{3.2}$$

In its analytical expansion the multiplicative model seems prohibitive compared to the additive model. However, it requires only the addition of a single parameter (*w*), which defines all interaction terms. Therefore, the type of interaction it models is rather constrained (von Winterfeldt and Edwards, 1986). Additive aggregation is the form that is most easily explained and understood by decision-makers from a wide variety of backgrounds, while not placing any substantially greater restrictions on the preference structures than more complicated aggregation formulae (Belton and Stewart, 2002).

In general the partial value functions should be standardised in a well-defined manner as will be described below. This is most easily done for criteria associated with measureable attributes, but it can be done quantitatively in other cases. Once an initial model structure like the above and a set of alternatives for evaluation have been defined, the next step will be to elicit the information required by the model. There are two types of information, sometimes referred to as intra-criterion information and inter-criteria information, or alternatively as scores and weights.

3.1.1 Eliciting scores

Scoring is the process of assessing a value derived by the decision-maker from the performance of alternatives against the relevant criteria. That is, the assessment of the partial value functions, $v_i(a)$ in the above model. If the criteria are structured as a value tree then the alternatives must be scored against each of the bottom level criteria. These values need to be assessed on an interval scale of measurement, i.e. a scale on which the difference between points is the important factor. A ratio of values will only have meaning if the zero point on the scale is absolutely and unambiguously defined. Thus to construct a scale it is necessary to define two reference points and to allocate numerical values to these points. The minimum and maximum points on the scale can be defined in a number of ways, e.g. 0 and 100, but it is important to distinguish between a local and a global scale:

> ***A local scale*** is defined by the set of alternatives that is under consideration. The alternative which does best on a particular criterion is assigned a score of 100 and the one that does least well is assigned a score of 0. All other alternatives will receive intermediate scores which reflect their performance relative to the end points. The use of local scales permits a relative quick assessment of values and can be very useful for initial "roughing out" of a problem, or if operating under time constraints.
>
> ***A global scale*** is defined by reference to the wider set of possibilities. The end points may be defined by the ideal and the worst conceivable performance on the particular criterion, or by the best and worst performance that can realistically occur. The definition of a global scale requires more work than a local scale. However, it has the advantages that it is more general than a local scale and that it can be defined before consideration of specific alternatives. This also means that it is possible to define criteria weights before consideration of alternatives.

Valid partial value functions can be based on either local or global scales. The important point is that all following analysis, including assessment of the weights (w_i), must be consistent with the chosen scaling. Once the reference points of the scale have been determined consideration must be given to how other scores are to be assessed. This can be done in one of the following three ways (Belton and Stewart, 2002):

1. **Definition of a partial value function.** This relates to performance in terms of a measurable attribute reflecting the criterion of interest.

2. **Construction of a qualitative value scale.** In this case, the performance of the alternatives can be assessed by reference to descriptive pointers, or word models to which appropriate values are assigned.

3. **Direct rating of the alternatives.** In this case, no attempt is made to define a scale which characterises performance independently of the alternatives being evaluated. The decision-maker simply specifies a number, or identifies the position on a visual analogue scale, which reflects the value of an alternative in relation to the specified reference points.

Definition of a partial value function

The first step in defining a value function is to identify a measurable attribute scale which is closely related to the decision-makers values. If it is not possible to identify an appropriate quantitative scale, or if such scales as are available are only remotely related to the decision-makers values then it will be necessary to construct a value scale (this will be described in the next section). The value function reflects the decision-makers' preferences for different levels of achievement on the measureable scale. Such a function can be assessed directly or by using indirect questioning. Direct assessment will often utilise a visual representation.

When making direct assessment of a value function the decision-maker should begin by determining whether:

- The value function is monotonically increasing against the natural scale, i.e. the highest value of the attribute is the most preferred and the lowest value the least preferred.
- The value function is monotonically decreasing against the natural scale, i.e. the lowest value of the attribute is the most preferred and the highest value the least preferred. This is e.g. the case with cost criteria.
- The value function is non-monotonic, i.e. an intermediate point on the scale defines the most preferred or least preferred point.

Von Winterfeldt and Edwards (1986) propose that if the value tree has been well structured then the value functions should be regular in form, i.e. no discontinuities. They go further to argue that all value functions should be linear or close to linear and propose that the analyst should consider restructuring a value tree to replace non-monotonic value functions by one or more monotonic functions. Whilst Belton and Stewart (2002) agree that an extremely non-linear value function, in particular a non-monotonic function, may indicate a need to revisit the definition of criteria, they caution against over-simplification of the problem by inappropriate use of linear value functions. Experimental simulations of Stewart (1993, 1996) propose that the results of analyses can be sensitive to such assumptions. Thereby, the default assumption of linearity, which is often made, may generate misleading answers.

Indirect assessment methods assume that the value function is monotonically increasing or decreasing over the range of attribute measurement considered. The end points of the scale must, as previously mentioned, be defined first. Thereafter, two methods of assessment are widely used, namely the bisection and the difference methods as described by von Winterfeldt and Edwards (1986).

Using the *Bisection method* the decision-maker is asked to identify the point on the attribute scale which is halfway, in value terms between the two endpoints. To help the decision-maker identify the midpoint value it may be helpful to begin by considering the midpoint on the objective scale and then pose a question regarding which of the two half's increase is the most valuable. The considered point can then be moved towards the most preferred half and the question repeated until the midpoint is identified. The next step would then be to find the midpoints between the two endpoints and the previous found midpoint. It is generally accepted that 5 points (2 endpoints and 3 "midpoints") give sufficient information to enable the analyst to sketch in the value function, see Figure 3.1.

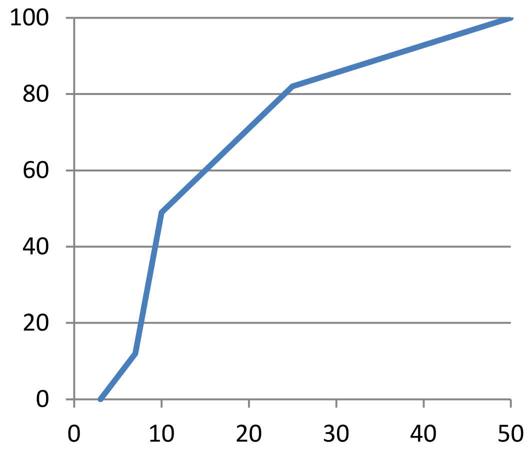


Figure 3.1. Example of a value function sketched using the bisection method

Difference methods could be viewed as a collection of methods rather than a single one, but all of them require the decision-maker to consider increments on the objectively measured scale and to relate these to difference in values. In the first approach as described by Watson and Buede (1987), the attribute scale is divided into, e.g., four equal intervals. To illustrate this approach, consider a simple example where a new department of a company has to hire people (besides the management) in order to make the department run. For the criterion "number of people" the minimum number is 0, and the maximum 36. Since preference is for more people, an increase in the number results in an increase in value. The decision-maker is asked to rank order the specified differences according to increase in associated value. For example, is the increase in value which occurs in going from 0 to 9 greater than, equal to or less than the increase in value achieved in going from 9 to 18? Suppose the information from the decision-maker is as given below in Table 3.1.

Table 3.1. Intervals on the criterion "number of people"

Increase in number of people		Increase in value
From	To	
0	9	1 = greatest
9	18	2
18	27	3
27	36	4

The ranking gives an idea of the shape of the value function. In this example the increase in value is greatest for low numbers of people, proposing a concave, increasing value function. The curve could be sketched directly on the basis of this information, as illustrated in Figure

3.2, or may be further refined by asking the decision-maker to assess the relative magnitude of value increases.

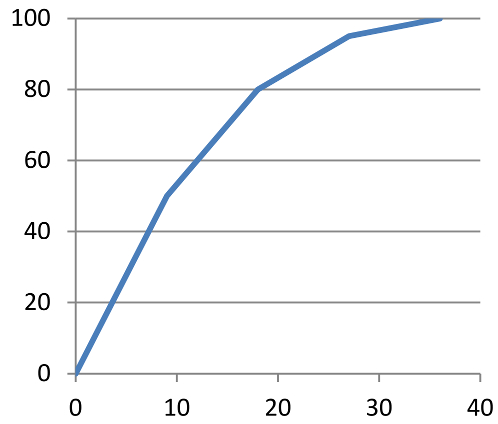


Figure 3.2. Example of a value function sketched using the difference method

Another approach is to begin by defining a unit level on the attribute scale (between one tenth and one fifth of the difference between the minimum and maximum points is proposed). Consider again the criterion “number of people”, measured as above. The minimum and maximum points on this scale is 0 and 36 people, thus let the specified unit be equal to 4 people (close to one tenth of the range). To assess the value function using this method we would first ask: What is the number of people, P, such that an increase from 4 to P people results in the same increase in value as an increase from 0 to 4 people? Suppose the decision-maker suggest that P should be 9. We next pose the question: What is the value of P such that an increase from 9 to P people is equal in value to the increase from 4 to 9? The decision-maker responds that it would be necessary to double the number of people in order to achieve the same increase in value. The additional value of 9 extra people then diminishes further, the increase from 18 to 36 perhaps equating in value to the increase from 9 to 18, but beyond 36 extra people do not add value. These responses give rise to a value function, specified in Table 3.2, which is very similar in shape to that defined using the previous method.

Table 3.2. Value function for the criterion “number of people”

Number of people	Value (per units defined above)	Value (0 to 100 scale)
0	0	0
4	1	25
9	2	50
18	3	75
36	4	100

The measurement scales used for the assessment of value functions for the two examples above arise naturally in the given context. Von Winterfeldt and Edwards (1986, p. 221) comment that:

> *"A natural scale which is linear in value is obviously the most economical device for communicating value relevant information"*

However, in some instances a simple natural scale may not exist and it becomes necessary to construct an appropriate measurement scale.

Construction of qualitative value scales

Often it is not possible to find a measureable attribute which captures a criterion. In such circumstances it is necessary to construct an appropriate qualitative scale. As discussed in the previous section, it is necessary to define at least two points on the scale (often taken as the end points). Intermediate points may also be defined. Points on the scale are defined descriptively and draw on multiple concepts in the definition (Belton and Stewart, 2002). An alternative approach for defining a scale could be to associate specific alternatives, with which the decision-makers are familiar, with points on the scale.

A qualitative scale should have the following characteristics (Ibid.):

- *Operational:* allow the decision-makers to rate alternatives not used in the definition of the scale
- *Reliable:* two independent ratings of an alternative should lead to the same score
- *Value relevant:* relates to the decision-makers' objective
- *Justifiable:* an independent observer could be convinced that the scale is reasonable

The approach described above directly assigns values to the qualitative statements. The MACBETH system by Bana e Costa and Vansnick (1994) can, as an example, be used to build a value scale from a category scale by a process of pair wise comparisons requesting ordinal assessments about preference differences. The output of the MACBETH system is a range of values associated with each category, consistent with the judgments input to the analysis. The decision-maker may choose to work with the midpoints of these intervals as the corresponding value scale, or may wish to further refine the input judgments to arrive at a tighter definition of values. It is possible that the initial judgments are ordinal inconsistent, in which case the method highlights inconsistencies and suggests revisions which would move towards consistency.

Direct rating

Direct rating can be viewed as the construction of a value scale, but defining only the end points of the scale. A local or global scale can be used, the former creating minimal work for the decision-makers. If using a local scale, the alternative which performs best of those under consideration is given the highest score, usually 100, and the alternative which performs least well (not necessarily badly in any absolute sense) is given a score of 0. All other alternatives are positioned directly on the scale to reflect their performance relative to the two end points. Although no attempt is made to relate performance to a measureable scale, the positioning of alternatives can generate extensive discussion, yielding rich information on the decision-makers' values. Ideally this information should be recorded for future reference. A disadvantage of using a local scale is that if new alternatives are introduced into the evaluation this may necessitate the revision of scales, something which has consequences for the weighting of criteria.

Direct rating by pair wise comparisons

The use of pair wise comparisons is implicit in all scoring processes as scores are assessed relative to reference points rather than in an absolute sense. Furthermore, in order to check consistency of judgments a facilitator may incorporate questioning processes which make explicit pair wise comparisons between alternatives. However, even if explicit, such comparisons tend to be ad-hoc and do not consider all possible comparisons. A systematic pair wise comparison approach is one of the cornerstones of the Analytic Hierarchy Process (AHP) by Saaty (1977, 2001). The AHP employs a process for direct rating which requires the decision-maker to consider all possible pairs of alternatives with respect to each criterion in turn, to determine which of the pair is preferred and to specify the strength of preference according to a semantic scale or the associated numeric 1-9 scale (see Section 3.3 for an elaboration of the technique). However, the AHP treats the responses as ratio judgments of preferences, which is not consistent with the value function approach. The underlying mathematics is easily modifiable to be consistent with difference measurement. The MACBETH approach mentioned before, which is founded on difference measurement and also based on pair wise comparisons, can be used to derive direct ratings. An additional approach, which can be used for the purpose of deriving direct ratings, is the REMBRANDT difference based technique by Lootsma (1992). The approach is also based on pair wise comparisons and overcomes some of the problems with the underlying mathematics of the AHP (this will be elaborated in Section 3.4).

One of the potential drawbacks of pair wise comparison methods is the large number of judgments required of the decision-maker: $n(n-1)/2$ for each criterion, where n is the number of alternatives. Nevertheless, the approach is powerful and can be effectively utilised if decision-makers find the direct rating process difficult. With some pair wise comparison approaches it is not necessary to compare all possible pairs and considerable work has been done to derive appropriate sampling processes (Belton and Stewart, 2002).

3.1.2 Eliciting weights

It is clear that in any evaluation not all criteria carry the same weight, thus it is desirable to incorporate an assessment of the relative importance of criteria. This aspect of analysis has been the focus of extensive debate (Belton and Stewart, 2002). Decision-makers are able and willing to respond to questions like: "what is most important to you when choosing a new car, safety or image?" Furthermore, they are able and willing to respond to questions asking them to rate the relative importance of safety and image against a numerical or verbal scale. The AHP technique is, as mentioned earlier, founded on such questions.

However, it has been argued that the responses to such questions are essentially meaningless. The questions are open to many different interpretations, people do not respond to them in a consistent manner and responses do not relate to the way in which weights are used in the synthesis of information (Ibid.). The weights which are used to reflect the relative importance of criteria in a multi-attribute value function are, however, well defined. The weight assigned to a criterion is essentially a scaling factor which relates scores on that criterion to scores on all other criteria. Thus if criterion A has a weight which is twice that of criterion B this should be interpreted as the decision-maker values 10 points on criterion A the same as 20 points on criterion B and would be willing to trade one for the other. These weights are often referred to as swing weights to distinguish them from the less well defined concept of importance weights. Thus the notion of swing weights captures both the psychological concept of “importance” and the extent to which the measurement scale adopted in practice discriminates between alternatives. One of the most common errors in naive scoring models is to assume that weights are independent of the measurement scales used. It can be seen from the algebraic structure of (3.1), however, that the effect of the weight parameter w_i is directly connected to the scaling used for $v_i(a)$, so that the two are intimately connected.

Swing weights

A method for eliciting weights for criteria is available with the swing weight technique (von WInterfeldt and Edwards, 1986). The technique is usually considered to be the theoretical most correct method for eliciting criteria weights, but also difficult to use in practice.

The technique presupposes that the decision-makers consider the swing from worst to best performance within each criterion and rank the criteria based on which swing gives the highest increase in overall value. Afterwards the swings within each of the criteria are assigned a numerical value reflecting its importance compared to the swing within the most important criterion.

It can be useful to work with graphically supported scales as decision-makers generally seem to be comfortable with this and may be willing to assess the relative magnitude of the swing weights directly using this means. An example of such graphical support is illustrated in Figure 3.3 where a small example concerning criteria for the selection of a new by-pass road is presented. First the criteria are ranked in order of importance by considering which swing from worst to best performance within each criterion that gives the greatest increase in overall value, the next greatest increase in overall value, and so on, until a ranking is established. The swing from worst to best within the highest ranked criterion is then assigned a value of 1 (see the column for “Urban development”). The swing from worst to best within the second highest ranked criterion (“Landscape”) is then using the visual scale compared with the swing within the highest ranked criterion. The process is repeated with the remaining criteria.

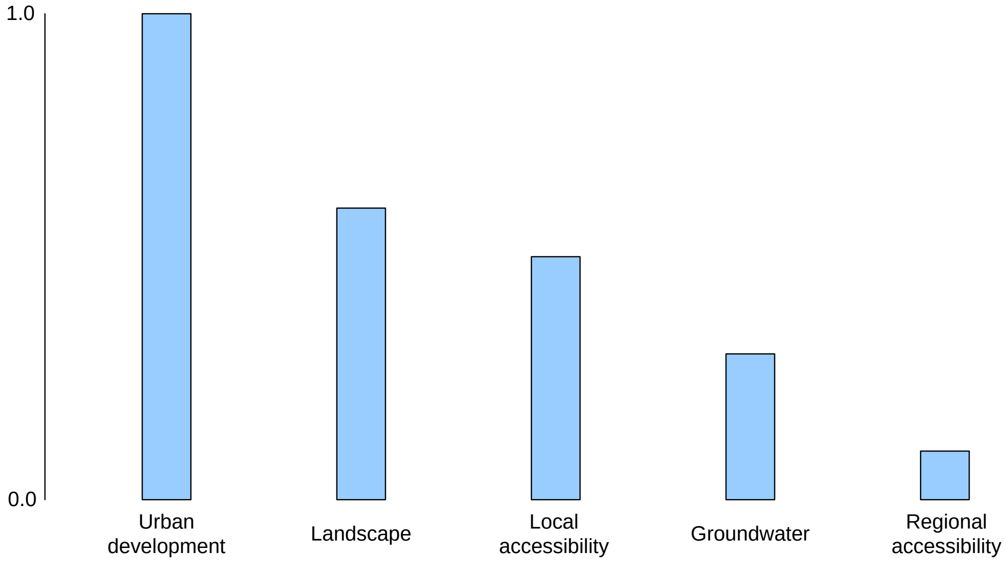


Figure 3.3. Swing weights: a visual analogue scale has been used for determining the magnitude between the five criteria in the example

This visual scaling provides a means for communicating a good sense of the magnitude of judgments whilst removing the need for numerical precision. However, it is important that this degree of imprecision is not forgotten when information is aggregated.

The weights implied by the visual representation in Figure 3.3 may be translated into numerical values, as shown in Table 3.3 below. The second column of the table lists the weights standardised with the largest weight set to 1. It is usual, although not essential, to normalise weights to sum to 1 or 100, as shown in the third column of Table 3.3. Such normalisation allows the decision-makers to interpret for example the weight of landscape in Table 3.3 as constituting 24% of the total importance weight. This often seems to be a useful interpretation. However, in specific cases decision-makers may find it more intuitive to specify a reference criterion whose units are weighted 1 and against which all other criteria are compared, as shown to be the original weights with urban development as the reference criterion.

Table 3.3. Swing weights – original and normalised values for the example

Criterion	Original weights	Normalised weights
Landscape	0.6	0.24
Groundwater	0.3	0.12
Urban development	1.0	0.40
Local accessibility	0.5	0.20
Regional accessibility	0.1	0.04

Weights in value trees

When the problem is structured as a multi-level value tree consideration has to be given to weights at different levels of the tree. It is useful to define relative weights and cumulative weights. Relative weights are assessed within families of criteria – i.e. criteria sharing the same parent – the weights within each family being normalised to sum to 1. The cumulative weight of a criterion is the product of its relative weight in comparison with its siblings and the relative weights of its parent, parent's parent, and so on to the top of the tree.

For illustration the example from above has been divided into different levels of criteria in a value tree, see Figure 3.4. By definition, the cumulative weights of all bottom-level criteria (leaves on the tree) sum to 1 – thus the normalised weights shown in Figure 3.4 are cumulative weights. The cumulative weight of a parent criterion is the total of the cumulative weights of its descendants.

As illustrated for the example problem, if the value tree does not have too many leaves, then weights can be assessed by directly comparing all bottom-level criteria to give the cumulative weights. Weights at higher levels of the tree are then to be determined by adding the cumulative weights of all members of a family to give the cumulative weight of the parent. Relative weights are determined by normalising the cumulative weights of family members to sum to 1. Relative and cumulative weights for the example problem are illustrated in Figure 3.4.

For larger models it is easier to begin by assessing relative weights within families of criteria. Weights at higher levels of the value tree can be assessed top-down or bottom-up. The top-down approach would assess relative weights within families of criteria by working from the top of the tree downwards. However, the analyst must be aware of the difficulty of interpreting weights at higher levels of a value tree – the weight of a higher level criterion is the sum of the cumulative weights of all its sub-criteria. Thus, in comparing two higher level criteria the decision-maker should be thinking in terms of a swing from 0 to 100 on all sub-criteria of the two higher level criteria. If the top-down approach is used it is important to carry out cross family checks on the cumulative weights of bottom level criteria.

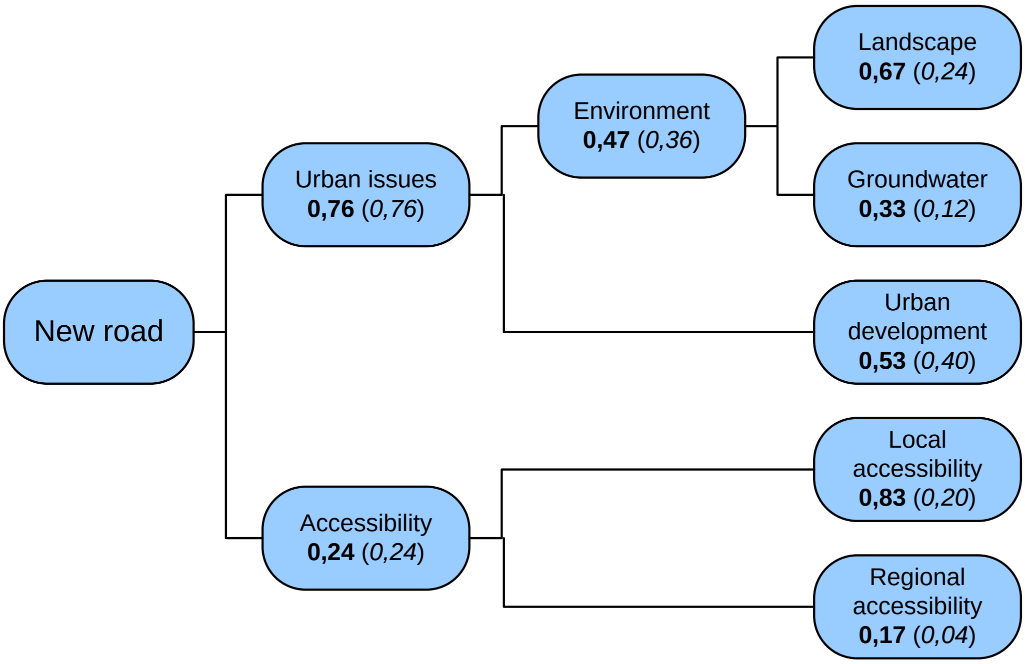


Figure 3.4. Relative weights (in bold) and cumulative weights (in italics) for the example.

The bottom-up approach begins by assessing relative weights within families which contains only bottom level criteria and then carrying out cross family comparisons using one criterion from each family (perhaps the most highly weighted criterion in each family) and comparisons with any unitary bottom level criteria. This process would eventually give the cumulative weights of the bottom level criteria which can be aggregated to higher levels as described before.

3.2 The SMART technique

The Simple Multi-Attribute Rating Technique (SMART) (von Winterfeldt and Edwards, 1986) is a method for MCDA whereby a finite number of decision alternatives are evaluated under a finite number of performance criteria. The purpose of the analysis is to rank the alternatives in a subjective order of preference and, if possible, to rate the overall performance of the alternatives via the proper assignment of numerical grades.

The SMART technique is based on a linear additive model. This means that an overall value of a given alternative is calculated as the total sum of the performance score (value) of each criterion multiplied with the weight of that criterion, see (3.1) in the first section of Chapter 3.

A SMART analysis includes the following five steps (modified from von Winterfeldt and Edwards (1986)):

Step 1: Define the alternatives and value-relevant criteria

Step 2: Evaluate each alternative separately on each criterion

Step 3: Assign relative weights to the criteria

Step 4: Aggregate the weights of criteria and the single-criterion evaluations of alternatives to obtain an overall evaluation of alternatives

Step 5: Perform sensitivity analysis and make recommendations

In SMART, ratings of alternatives are assigned directly, in the natural scales of the criteria. For instance, when assessing the criterion "cost" for the choice between different road layouts, a natural scale would be a range between the most expensive and the cheapest road layout. In order to keep the weighting of the criteria and the rating of the alternatives as separate as possible, the different scales of criteria need to be converted into a common internal scale. In SMART, this is done mathematically by the decision-maker by means of a value function. As mentioned previously the simplest and most widely used form of a value function method is the additive model, which in the simplest cases can be applied using a linear scale (e.g. going from 0 to 100).

3.2.1 SMART Exploiting Ranks (SMARTER)

The assessment of value functions and swing weights in SMART can sometimes be a difficult task, and decision-makers may not always be confident about it. Because of this, Edwards and Barron (1994) have proposed a simplified form of SMART named SMARTER (SMART Exploiting Ranks). Using the SMARTER technique the decision-makers place the criteria into an importance order: for example “Criterion 1 is more important than Criterion 2, which is more important than Criterion 3, which is more important than Criterion 4" and so on, $C_1 \geq C_2 \geq C_3 \geq C_4$. . . . SMARTER then assigns "surrogate" weights to the criteria.

Barron and Barret (1996) believe that generated weights may be more precise than weights produced by the decision-makers who may be more comfortable and confident with a simple ranking of the importance of each criterion swing, especially if it represents the considered outcome of a group of decision-makers. Therefore a number of methods that enable the ranking to be translated into "surrogate" weights representing an approximation of the "true" weights have been developed. These are among others Rank Order Centroid (ROC), Rank Sum (RS), RR Rank Reciprocal (RR) and Rank Order Distribution (ROD) weights. Roberts and Goodwin (2002) have examined these methods in details and found that ROD weights seem to provide the best approximation to decision-makers preferences.

ROD is a weight approximation method that assumes that valid weights can be elicited through direct rating. In the direct rating method the most important criterion is assigned a weight of 100 and the importance of the other criteria is then assessed relative to this benchmark. The "raw" weights, (w_i^*) obtained are then normalised to sum to 1. Assuming that all criteria have some importance, this means that the ranges of the possible "raw" weights will be:

$$w_1^* = 100, \qquad 0 < w_2^* \leq 100, \quad 0 < w_3^* \leq w_2^*$$

And in general:

$$0 < w_i^* \leq w_{i-1}^* \text{ (where } i \neq 1)$$

These ranges can be approximated by representing all of the inequalities by less-than-or-equal-to expressions. The uncertainty about the "true" weights can then be represented by assuming uniform distribution for them. To determine ROD weights for general problems it is needed to consider the probability distributions for the normalised weights that follow from the assumptions about the distributions of the "raw" weights. For $n > 2$ the density functions are a series of piecewise equations.

Compared with other surrogate weight approximation methods, as mentioned before, the use of ROD weights goes some way to reduce the extreme value problem of having criteria with very low weights in the assessment. However, it can be argued that the inclusion of criteria with very low weights, e.g. 0.02, does not contribute in any way to the overall result and therefore should be omitted from the analysis. A discussion of this issue can be found in Paper 2 in this thesis.

The means of each rank order distribution for n = 2 to 10 have been found mathematically and are displayed in Table 3.4. For further information about the calculations behind see Roberts and Goodwin (2002).

Table 3.4. Rank Order Distribution (ROD) weights (Roberts and Goodwin, 2002)

	Attributes								
Rank	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	0.6932	0.5232	0.4180	0.3471	0.2966	0.2590	0.2292	0.2058	0.1867
2	0.3068	0.3240	0.2986	0.2686	0.2410	0.2174	0.1977	0.1808	0.1667
3		0.1528	0.1912	0.1955	0.1884	0.1781	0.1672	0.1565	0.1466
4			0.0922	0.1269	0.1387	0.1406	0.1375	0.1332	0.1271
5				0.0619	0.0908	0.1038	0.1084	0.1095	0.1081
6					0.0445	0.0679	0.0805	0.0867	0.0893
7						0.0334	0.0531	0.0644	0.0709
8							0.0263	0.0425	0.0527
9								0.0211	0.0349
10									0.0173

It should be noted that the four decimals that is shown for the ROD weights in Table 3.4 express a much higher accuracy in the weights than should be expected in practice. Normally, decision-makers assign weights with not more than two decimals as this seems to be the limit to what can be comprehended by the human mind without difficulties. Thus the weights in Table 3.4 should be presented with only two decimals to the decision-makers if this technique is used in the decision process.

3.2.2 Strengths and weaknesses of SMART

The structure of the SMART method is similar to that of the conventional cost-benefit analysis (CBA) in the sense that the total value is calculated as a weighted sum of the impact scores. In the CBA the unit prices act as weights and the impacts scores are the quantified (not normalised) CBA impacts. This close relationship to the well-accepted CBA method is appealing and makes the method easier to grasp for the decision-makers.

However, in a screening phase where some poorly performing alternatives are rejected leaving a subset of alternatives to be considered in more detail the SMART method is not always the right choice. This is because, as noted by Hobbs and Meier (2000), SMART tends to oversimplify the problem if used as a screening method as the top few alternatives are often very similar. Rather different weight profiles should be used and alternatives that perform well under each different weight profile should be picked out for further analysis. This also helps to identify the most robust alternatives. The SMART method has rather high demands on the level of detail in input data. Value functions need to be assessed for each of the lowest-level criteria, and weights should be given as trade-off.

In a SMART analysis the direct rating method of selecting “raw” weights is normally used as it is cognitively simpler and therefore is assumed to yield more consistent and accurate judgments from the decision-maker. These “raw” weights are then normalised and this normalisation process yields different theoretical distributions for the ranks. The means of these distributions are the ROD weights.

The formula for the distribution of the ROD weights becomes progressively more complex as the number of criteria increases. Since RS weights are easy to calculate and closely match the ROD weights for higher numbers of criteria Roberts and Goodwin (2002) recommend using RS weights when working with problems involving large numbers of criteria, and in cases where it can be assumed that the appropriate alternative method for eliciting the "true" weights would have been the direct rating method.

3.3 The Analytic Hierarchy Process

The Analytic Hierarchy Process (AHP), developed by Saaty (1977), is essentially the formalisation of our intuitive understanding of a complex problem using a hierarchical structure (Hwang and Yoon, 1995). The AHP offers an alternative approach to SMART when a decision-maker is faced with a problem involving multiple objectives.

The crux of the AHP is to enable a decision-maker to structure a decision problem visually in form of an attribute hierarchy. An attribute hierarchy has at least three levels: the focus or the overall goal of the problem on the top level, multiple criteria that define the alternatives in the middle level, and the competing alternatives in the bottom level as depicted on Figure 3.5. When criteria are highly abstract such as e.g. “landscape”, sub-criteria (or sub-sub-criteria) are generated subsequently through a multilevel hierarchy.

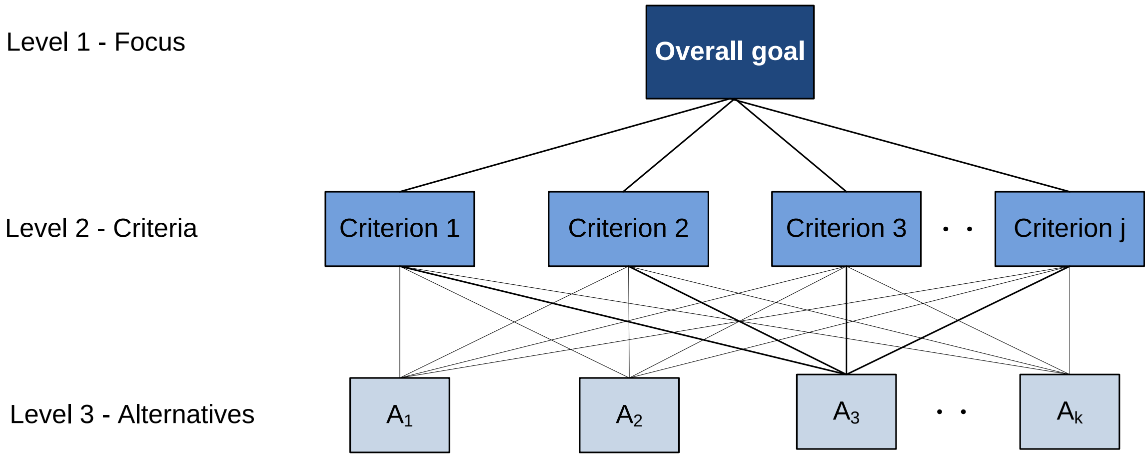


Figure 3.5. An attribute hierarchy

Each criterion in level 2 will possibly contribute differently to the focus. The decision can be made on the relative importance among four criteria by pair wise comparisons, due to the fact that pair wise comparisons are much easier to make than a comparison of four criteria simultaneously.

In order to help the decision-maker to assess the pair wise comparisons, Saaty created a nine point intensity scale of importance between two elements (Saaty, 1977; 2001). The verbal scale and the proposed numbers to express the degree of preference between the two elements are shown in Table 3.5.

Table 3.5. The fundamental scale for pair wise comparisons (Saaty, 2001)

Intensity of importance	Definition	Explanation
1	Equal importance	Two activities contribute equally to the objective
3	Moderate importance	Experience and judgment slightly favour one activity over another
5	Strong importance	Experience and judgment strongly favour one activity over another
7	Very strong or demonstrated importance	An activity is favoured very strongly over another; its dominance demonstrated in practice
9	Extreme importance	The evidence of favouring one activity over another is of the highest possible order of affirmation
2, 4, 6, 8	For compromise between the above values	Sometimes one needs to interpolate a compromise judgment numerically because there is no good word to describe it
Reciprocals of above	If activity k has one of the above nonzero numbers assigned to it when compared with activity j, then j has the reciprocal value when compared with k	A comparison mandated by choosing the smaller element as the unit to estimate the larger one as a multiple of that unit

The technique interprets the above numerical scale of strengths of preferences in a ratio sense which means that the decision-makers are allowed to quantify and compare the sizes of ratios between the alternatives. For example, the Kelvin temperature scale is a ratio scale, not only can it be stated that a temperature of 200 degrees is higher than one of 100 degrees, it can correctly be stated that it is twice as high. Interval scales such as e.g. the Celcius scale do not have the ratio property.

It should be noted that the verbal statements on Saaty's original scale in Table 3.5 have been modified by other researchers to go from indifference, weak, definite, strong to very strong preference for one object over another. The numerical scale, however, remains unchanged in this original additive version of AHP.

To decide the relative weightings between *n* alternatives, it is in principle only necessary to perform *n-1* assessments. By performing a complete set of full pair wise comparisons more information than necessary is collected, but a more varied evaluation is obtained, and if one or more answers are inaccurate the other answers will compensate the inaccuracy. The number of judgments, *J*, that have to be made in a full pair wise comparison can be determined by (Belton and Stewart, 2002):

$$J = \frac{n \cdot (n-1)}{2} \tag{3.3}$$

The information from the pair wise comparisons can be concisely contained in a so-called comparison matrix whose element at row *k* and column *j* is the ratio of row *k* and column *j* (Hwang and Yoon, 1995). The comparison matrix *A*, as introduced by Saaty, is seen below:

$$A = \begin{bmatrix} \frac{w_1}{w_1} & \cdots & \frac{w_1}{w_n} \\ \vdots & \ddots & \vdots \\ \frac{w_n}{w_1} & \cdots & \frac{w_n}{w_n} \end{bmatrix} \tag{3.4}$$

Where $w_1, \ldots, w_n$ is the weights obtained by the comparisons. After the construction of the pair wise comparison matrix, the next step is to retrieve the scores of each element in the matrix. There are basically two different methods for retrieving these scores: the originally introduced eigenvector method (Hwang and Yoon, 1981), and the later introduced geometric mean method (Barzilai et al., 1987).

3.3.1 Consistency

The AHP allows inconsistency, but provides a measure of the consistency in each set of judgments. This measure is an important by-product of the process of deriving priorities based on pair wise comparisons.

It is important that a low consistency ratio (CR) does not become the goal of the decision making process. A low CR is necessary but not sufficient for a good decision. It is possible to be perfectly consistent but consistently wrong. It is more important to be accurate than consistent.

The CR is computed from the eigenvalue, λ_{max}, which will often turn out to be larger than the value describing a fully consistent matrix. In order to provide a measure of severity of this deviation, Saaty defined a measure of consistency, or consistency index (CI) by:

$$CI = \frac{principal\ eigenvalue - size\ of\ matrix}{size\ of\ matrix - 1} = \frac{\lambda_{max} - n}{n-1} \tag{3.5}$$

The CI is compared to a value derived by generating random reciprocal matrices of the same size, to give a CR which is meant to have the same interpretation no matter the size of the matrix. The comparative values (CV) from random matrices are as follows in Table 3.6 for 3 ≤ n ≤ 10 (Belton and Stewart, 2002, p. 156):

Table 3.6. Comparative values (CV)

Size of matrix	***3***	***4***	***5***	***6***	***7***	***8***	***9***	***10***
Comparative value	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49

Hence, the CR is calculated as follows:

$$CR = \frac{CI}{CV} \tag{3.6}$$

A CR of 0.1 or less is generally stated to be acceptable (Ibid.). If the CR is higher there will be a need of revising some of the judgments in the comparison matrix.

3.3.2 Strengths and weaknesses of the AHP

This section highlights the strengths of the AHP, but several criticisms which have been made of the technique are also presented. These are mainly based on the work of Stewart (1992, 1993) and Goodwin and Wright (2009).

In short the strengths of the technique are:

- AHP provides a formal structure to problems. This allows complex problems to be decomposed into sets of simpler judgments and provides a documented rationale for the choice of a particular option.
- Only two attributes or alternatives have to be considered at any one time so that the decision-maker's judgmental task is simplified.
- It is considered to be good practice in decision analysis to obtain an input for a decision model by asking for it in several ways (redundancy) and then asking the decision-maker to reflect on any inconsistencies in the judgments put forward.
- The wide range of applications of the AHP is evidence of its versatility; see e.g. Vaidya and Kumar (2006).

The weaknesses can be summed to the following:

- The correspondence between the verbal and the numerical scales is based on untested assumptions.
- In some problem situations the restriction of pair wise comparisons to a 1 to 9 scale is bound to force inconsistencies on the decision-maker.
- Unlike SMART, weights are elicited in the AHP without reference to the scales on which attributes are measured. AHP questions, which simply ask for the relative importance of attributes without reference to their scales, are therefore less well defined, if they are meaningful at all. This fuzziness may mean that the questions are interpreted in different, and possibly erroneous, ways by the decision-makers.
- New alternatives can reverse the rank of existing alternatives. Goodwin and Wright (2009) claims that this arises from the way in which the AHP normalises the weights to

sum to 1, and that this is consistent with a definition of weights which is at variance with that used in SMART.

- While the redundancy built into the AHP is an advantage, it may also require a large number of judgments from the decision-maker. This requirement to answer a large number of questions can reduce the attraction of the AHP in the eyes of potential users, even though the questions themselves are considered to be easy.

- The axioms of the AHP are not founded on testable descriptions of rational behaviour.

The method has also attracted much controversy from people who have questioned its underlying axioms and the extent to which the questions that it poses can lead to meaningful responses from decision-makers. Indeed, it has been argued that the apparent simplicity of the questions belies a lack of clarity in their definition and may lead to superficial and erroneous judgments. Critics have also questioned the extent to which an AHP model can faithfully represent a decision-maker's preferences given the numerical representations of these judgments and the mathematical processes which are applied to them.

It should, however, not be forgotten that the purpose of any decision aid is to provide insights and understanding, rather than to prescribe a "correct" solution. Often the process of attempting to structure the problem is more useful in achieving these aims than the numeric output of the model. Nevertheless this process is still best served when the analytic method poses unambiguous questions and bases its proposed solutions on testable axioms and an accurate translation of the decision-maker's judgments (Goodwin and Wright, 2009). Whether the AHP is the best technique to support this process is a question which is bound to continue to attract debate and controversy. The next section presents an alternative version of the AHP which proposes to overcome some of the weaknesses described here.

3.4 The REMBRANDT technique

The REMBRANDT (Ratio Estimations in Magnitudes or deci-Bells to Rate Alternatives which are Non-DominaTed) technique is developed in order to adjust for three main weaknesses in the AHP technique. First, the direct rating in REMBRANDT is on a geometric scale (Lootsma, 1988) which replaces Saaty's 1 – 9 ratio scale. Second, the eigenvector method originally used in AHP is replaced by the geometric mean, which avoids potential rank reversal (Barzilai et al., 1987). Third, the aggregation of scores by arithmetic mean is replaced by the product of alternative relative scores weighted by the power of weights obtained from analysis of hierarchical elements above the alternatives (Olson, 1996).

As in the original AHP the decision-makers' pair wise comparative judgment of alternative A_j versus alternative A_k is in REMBRANDT captured on a category scale to frame the range of possible verbal responses. This is converted into an integer-valued gradation index δ_{jk} according to the scale in Table 3.7.

Table 3.7. The REMBRANDT scale (van den Honert and Lootsma, 2000)

Comparative judgment	Gradation index δ_{jk}
Very strong preference for A_k over A_j	-8
Strong preference for A_k over A_j	-6
Definite preference for A_k over A_j	-4
Weak preference for A_k over A_j	-2
Indifference	0
Weak preference for A_j over A_k	+2
Definite preference for A_j over A_k	+4
Strong preference for A_j over A_k	+6
Very strong preference for A_j over A_k	+8

Intermediate integer values can be assigned to δ_{jk} to express a hesitation between two adjacent categories. The gradation index δ_{jk} can be converted into a value on a geometric scale, characterised by a scale parameter $\gamma = \ln(1 + \varepsilon)$, where $1 + \varepsilon$ is the progression factor. Thus

$$r_{jk} = \exp(\gamma\delta_{jk}), \qquad j, k = 1, ..., n \tag{3.7}$$

is defined to be the numeric estimate of the preference ratio. Although there is no unique scale of human judgment, a plausible value of γ is $\ln(2)$ implying a geometric scale with the progression factor 2 (Lootsma, 1992). The progression factor 2 is closely related to human consistencies when categorising certain intervals of interest in totally unrelated areas, e.g. how human subjects partition certain ranges on the time axis and how they categorise sound and light intensities (Lootsma, 1999). Analysing such ranges Lootsma found geometric sequences with the progression factor 4 between basic categories describing a certain development. The number of categories is usually rather small as human beings' linguistic capacity to describe the categories unambiguously in verbal terms is limited. Thus there are five major, linguistically distinct categories in Table 3.7: indifference, weak, definite, strong and very strong. Moreover, there are four so-called threshold categories between them which can be used if the decision-makers are in-between the neighbouring qualifications. By the interpolation of threshold categories a more refined subdivision of the given interval is obtained. In that case there are nine categories and the progression factor is roughly 2 (Lootsma, 1993). When determining criteria weights Lootsma (1999) finds the progression factor to be √2.

The second improvement is the calculation of impact scores. The arithmetic mean is subject to rank reversal of alternatives. The geometric mean is not subject to rank reversal, nor is logarithmic regression. Barzilai et al. (1987) have argued that the geometric mean is more appropriate for calculation of relative value through weights than the arithmetic mean used by Saaty.

Lootsma proposes logarithmic regression, minimising:

$$\sum_{j<k} \left(\ln r_{jk} - \ln v_j + \ln v_k\right)^2 \tag{3.8}$$

Where r_{jk} are the ratio comparisons made by the decision-maker for object j and compared object k. The weight for j (w_j) is represented by $ln\ v_j$. The ratio r_{jk} is the ratio of w_j/w_k. The analysis is then to calculate these weights. Since $r_{jk} = w_j/w_k$, error is represented by $r_{jk} - w_j/w_k$. The ratio comparisons made by the decision-maker are observations, and regression minimising the squared error yields the set of weights w_i which best fit the decision-maker expressed preferences. Solving this is complicated by the fact that the resulting data set is singular. However, a series of normal equations can be solved to yield the desired weights (Olson, 1996).

The third improvement constructed by Lootsma is the aggregation of scores. The lowest level in the hierarchical structure of the decision problem, the alternative level, is normalised multiplicatively, so that the product of components equals 1 for each of the k factors over which the alternatives are compared. Therefore, each alternative has an estimated relative performance w_k for each of the k factors. The components of the hierarchical level immediately superior to this lowest level are normalised additively, so that they add to 1, yielding weights $O(j)$. The aggregation rule for each alternative j is (Olson et al., 1995):

$$w_j = \prod_{i=1,k} w_i^{O(i)} \tag{3.9}$$

Hence, the REMBRANDT technique is based on the multiplicative value function where the original AHP is based on the additive value function. For this reason reference is often made to the additive AHP when dealing with the original version and the multiplicative AHP when dealing with the REMBRANDT technique.

3.5 The composite model for assessment

In the appraisal and planning of transport infrastructure projects the examination should be based on all relevant impacts, which are depending on the type and size of the project viewed upon. Some of these impacts can be assessed monetarily and are thereby possible to include in a conventional CBA. However, no valid monetary assessment knowledge exists for impacts such as urban development, landscape, etc. These impacts are denominated as non-monetary impacts or strategic impacts and have to be assessed by use of a MCDA.

The idea behind composite modelling assessment (COSIMA) is to extend conventional CBA into a more comprehensive type of analysis, as often demanded by decision-makers, by including “missing” decision criteria of relevance for the actual assessment task. The “missing” criteria often address issues that have been difficult to assess by the conventional CBA but hold a potential of improving actual decision support from the assessment if treated properly. In COSIMA the added criteria will be referred to as the MCDA part of the COSIMA analysis.

The Danish manual for socio-economic analysis (DMT, 2003) describes the need for taking non-monetary impacts into account when assessing larger infrastructure projects. However, no specific guidelines are set out for how to include these impacts. A suggestion by the manual is to describe the impacts in an overview table and keep them in mind during the decision making. A more traditional way, however, of combining CBA and MCDA is to translate the CBA result into a part of the MCDA by applying scores to e.g. the B/C-rates. However, these methods “hide” the economic argument within the combined analysis which is considered to be a big weakness seen from a decision-makers point of view. For this reason the COSIMA methodology is set out based on keeping the economic information intact at all times.

3.5.1 Principles for composite modelling assessment

The COSIMA approach consists of a CBA part and a MCDA part and the result of the COSIMA assessment is expressed as a total value (TV) based on both parts. This model set-up emphasises that the MCDA part should be truly additive to the CBA part. For this reason a project alternative, A_k, is better represented for the decision making by the $TV(A_k)$ than by e.g. the net present value (NPV) derived from the CBA. The principle in COSIMA can be expressed by (Leleur et al., 2007):

$$TV(A_k) = CBA(A_k) + MCDA(A_k) \tag{3.10}$$

The formulation of COSIMA introduced by the equation thus resembles CBA but the assessment principles made use of in the MCDA part, generally based on decision-maker involvement, justifies the notation as MCDA. It can be noted on the basis of (3.10) that in a situation where the investment in A_k equal to the investment costs C_k is not feasible seen from CBA (i.e. $CBA(A_k) \leq C_k$), then the investment can be justified by the wider COSIMA examination if $TV(A_k) > C_k$. This can also be expressed as $TRR(A_k) > 1$ where TRR expresses the total rate of return.

In a COSIMA analysis where A_k denominates the project alternative it has been found to be convenient to express the feasibility by the total rate of return $TRR(A_k)$ from the investment C_k which leads to (Ibid.):

$$TRR(A_k) = \frac{TV(A_k)}{C_k} = \frac{1}{C_k} \cdot \left(\sum_{i=1}^{I} V_i(X_{ik}) + \alpha \cdot \left[\sum_{j=1}^{J} w_j \cdot VF_j(Y_{jk}) \right] \right) \tag{3.11}$$

where

$$\sum_{i=1}^{I} w_j = 1 \text{ and } 0 < w_j < 1$$

A_k is alternative k

C_k are the total costs or expenses of alternative k

X_{ik} is the quantity of CBA impact i for alternative k

$V_i(X_{ik})$ is the value in monetary units for the CBA impact i for alternative k

α is a indicator that expresses the model set-up's trade-off between the CBA and the MCDA part

w_j is a importance weight for criterion j

Y_k is a parameter value for MCDA criterion j for alternative k

$VF_j(Y_{jk})$ is a value-function score for MCDA criterion j for alternative k

The general COSIMA principles are presented by (3.10) and (3.11). It can be realised that with sufficient information about the MCDA part, the first equation can be specified into a CBA. This would for example be the situation if a conventional CBA is carried out and afterwards supplemented with some extra criteria which can be specified fully by impact models that lead to net effects which can be given satisfactory unit prices similar to the assessment in the CBA part. However, this will most often not be possible as the MCDA part in general is "less known" than the CBA part. The purpose of COSIMA is to handle such a situation in a comprehensive and transparent way ensured through the determination of appropriate values for α and w_j for the J MCDA criteria and appropriate value function scores $VF_j(Y_{jk})$. $V_i(X_{ik})$ can be derived from a CBA manual relevant for the actual assessment case e.g. DMT (2003).

3.6 Decision conferences

The decision-maker may be a single individual, a small group of individuals with more or less common goals, or e.g. a corporate executive or a political decision-maker acting on behalf of a large group of interested and affected persons with divergent interests. For a single individual, or a small homogeneous group with shared objectives, the decision making process can be used to identify the final decision directly, without the need to justify or to debate this with other groups.

For decision making involving groups with more divergent objectives, the final decision is likely to involve some form of political negotiation between stakeholders, each of whom may adopt different sets of criteria for evaluating alternatives. The final decision-maker will need to take all of these criteria into account in seeking a political consensus or compromise. The analysis may need to be conducted within a group setting involving representatives of all stakeholders, or may be carried out separately for sub-groups as a form of scoping exercise or impact assessment. For the purpose of handling these issues a decision making process – the so-called decision conference – is set out below.

Decision conferences as a concept was introduced in USA back in 1979 and the first European decision conference was held at Brunel University in England in 1981 (Goodwin and Wright, 2009). A decision conference joins the concepts of decision analysis, group processes and information technology in an intensive session where different persons involved in the decision process are present. A decision conference can be defined as the common area in Figure 3.6.

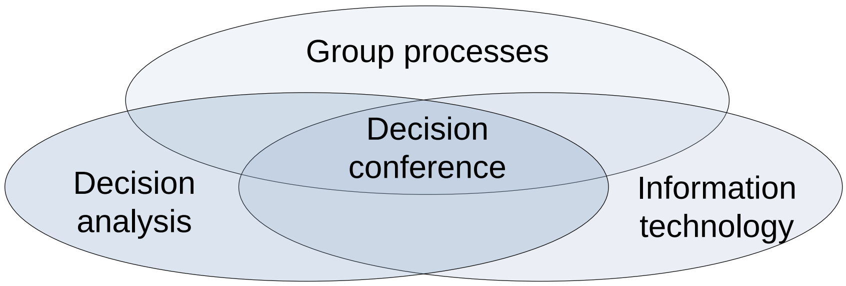


Figure 3.6. The joining of concepts at a decision conference

In the ideal situation, a decision conference which has the purpose of solving important issues is (Jeppesen, 2009; Goodwin and Wright, 2009):

- A meeting which in time can vary between half a day and three days dependent on the complexity of the decision problem, the number of participants involved, the number of alternative solutions, the number of criteria, etc.
- Attended by key persons which represent different perspectives on the decision problem

- Facilitated by an impartial facilitator
- Supported by a decision model which is operated by a decision analyst

A decision conference is an approach that makes it possible for a group of stakeholders and decision-makers representing very different viewpoints to work together in a way so efficient that they can create a vision based decision with regard to the common objective.

The decision conference can for instance take place by a group of decision-makers is being placed around a table with the purpose of discussing the issue. The conference is, as mentioned above, controlled by a facilitator who organises and facilitates the interplay and knowledge sharing in the group. In the background a decision analyst uses interactive decision support technology to model the issues and viewpoints which appear during the process.

It is important for the quality of the decision conference that the facilitator starts by introducing the underlying theory of the decision model for the decision-makers. This contributes to make the decision-makers more comfortable with the later decisions when they know how the model works. The decision process is, however, built up in such an intuitive way that the participants do not need a thorough knowledge of the theories and techniques applied.

The fundamental objective of a decision conference is to create a synthesis of decision analytical techniques and the positive features and dynamics which are found by decision making in smaller groups. Common understanding of the issues is created by decision techniques and social interaction. Hereby, the participants obtain a sense of the common objective and obligate themselves to act with a view to implementation. Moreover, sensitivity analysis gives the participants a possibility to see if individual disagreements make a decisive difference towards the final decision.

According to Phillips (2007) a decision conference is not about providing the best answer, it is about providing insight. The decision conference helps to conduct the MCDA according to the preferences of the participants in a comprehensive and transparent way. Decisions made in consensus at a decision conference have a fairly higher probability for being implemented than results from a complex decision analysis that only involve one decision-maker which later has to justify his decision for other people in the organisation or to the public. Moreover, a decision made by such a group has better terms for working in practise as it has the group’s commitment.

However, there is one large question that has to be answered: Are decisions made in consensus at a decision conference more or less valid than assessments and solutions made without aids? According to Phillips (2007) this is not necessarily the case, but a decision conference has the following advantages:

- The participants are not on home ground. Often decision conferences take place in hotels or in a specially designed room on the decision analyst’s premises.

- The group is carefully composed of people representing all perspectives on the issue to be resolved so that adversarial processes operate in the group to check bias and explore alternative framings of the decision problem.

- The decision analyst who acts to facilitate the conference is a neutral outsider who is sensitive to the obstructive effects of groupthink and reflects this back to the group.

The above mentioned has been reformulated by McCartt and Rohrbough (1989) to concern an appraisal of the effectiveness of a decision conference. The researchers claim that attempts to unite good decision results with certain types of decisions made by groups is difficult as practically all use in practice do not provide a satisfactory standard of comparison to satisfy proper research. However, Chun (1992), whose surveys are based on interviews with various companies which have used decision conferences, has found the result that a majority of the participants in 48 different decision conferences preferred this to ordinary meetings. More recently the usefulness of the concept has been proved by Mustajoki et al. (2004).

The concept of decision conferences is more thoroughly described and treated with regard to the preliminary planning in Jeppesen (2009). In paper 3 of this thesis the concept is operationalised in the context of a real decision problem.

4 Case examination and findings

The following chapter is a review of the work conducted in the field of decision support systems and multi-criteria decision analysis (MCDA) under this Ph.D. thesis. The work has ultimately resulted in a number of accepted peer-reviewed papers in international journals or conference proceedings.

The chapter comprises different aspects of the work, where the emphasis in the early work is on the combination between cost-benefit analysis (CBA) and MCDA. Furthermore, some considerations in the early papers (paper 1-2) are also made regarding the decision process and decision-maker interaction. These considerations are further explored in the papers 3 and 4, where guidelines for conducting decision conferences are set out and the structuring process of a decision problem is explored. The key references in these two papers mainly rely on von Winterfeldt and Edwards (1986, 2007), Phillips (1984, 2007), Belton and Stewart (2002), Keeney and Raiffa (1993) and Keeney (1992). The final and more recent paper (5) conducts a sensitivity examination of the scaling issues in the MCDA technique REMBRANDT, which is applied as assessment technique in the papers 2-5.

The five papers are all case specific where different infrastructure proposals are investigated using the applied methodologies. The sequence in which the papers appear is based on the progress of the decision support framework. The cases can be divided into three types of travel modes: road, railway and bikes. More specifically the five cases can be divided into respectively a fixed link, two railway cases, bikes and finally another fixed link. A schematic overview of the papers together with their main purposes is listed below.

1. ***Composite decision support by combining cost-benefit and multi-criteria decision analysis***
   This paper concerns a composite assessment of alternatives for a new fixed link connecting the peninsula of Hornsherred to the city of Frederikssund in Denmark using the following methodologies: Cost-benefit analysis and multi-criteria decision analysis.

2. ***Examination of decision support systems for composite CBA and MCDA assessments of transport infrastructure projects***
   This paper compares two decision support systems for the assessment of alternatives for a new high-speed railway line in Sweden using the following methodologies: Cost-benefit analysis and multi-criteria decision analysis.

3. ***Customised DSS and decision conferences***
   This paper investigates the decision process for the appraisal of alternatives for a new high-speed railway line in Sweden using the following methodologies: Multi-criteria decision analysis and decision conferences

4. ***An MCDA approach for the selection of bike projects based on structuring and appraising activities***
   This paper examines the decision structuring process related to the prioritisation task of selecting bike projects from a public pool with limited funds using the following methodologies: Decision analysis and multi-criteria decision analysis.

5. ***Scaling transformation in the REMBRANDT technique: a sensitivity examination of the progression factors***
   This paper examines the sensitivity related to the scaling issues in the MCDA technique REMBRANDT. The issues are illuminated using a case study for a new fixed link between Denmark and Sweden.

In each of the following sections, a small description of the case is given followed by the specific findings from the associated paper. The full papers are included in the end of this thesis.

4.1 Paper 1

Composite decision support by combining cost-benefit and multi-criteria decision analysis

Author(s): Barfod, M.B., Salling, K.B. and Leleur, S.

Published in *Decision Support Systems* 51, Issue 1, pp. 167-175, Elsevier 2011

Description

The paper presents the COSIMA decision support system (DSS) based on combining cost-benefit analysis (CBA) with multi-criteria decision analysis (MCDA) for the assessment of economic as well as strategic impacts of transport projects. The DSS ensures that the assessment is conducted in a systematic, transparent and explicit way. The modelling principles presented are illuminated with the assessment task of finding the most attractive alternative for a new fixed link between the city of Frederikssund and the peninsula of Hornsherred in Denmark. In this respect four different alternatives are examined, namely a high-level bridge, a short tunnel, a long tunnel, and an upgrade of the already existing connection.

The COSIMA DSS provides a theoretical and practical methodology for adding non-monetary MCDA-criteria to the monetary CBA-impacts. Unlike previous attempts the DSS is based on the argument that the MCDA-criteria can be added to the CBA-impacts – if value functions can be computed for the MCDA-criteria using a weighting process describing the importance of each criterion. Hence, the COSIMA approach is based on the theoretical valid and widely used methodology of additive value functions. The value function scores are in the presented COSIMA set-up derived via direct rating using pair wise comparisons and the criteria weights are derived by applying rankings in this respect drawing on the Analytic Hierarchy Process (AHP) and the Simple Multi-Attribute Rating Technique Exploiting Ranks (SMARTER) respectively. The total benefits from the CBA are used to determine shadow prices for the MCDA-criteria; this is based on a trade-off between the CBA and MCDA, which creates a total rate of return (TRR). The last part requires converting strategic non-quantitative criteria into monetary values. This is discussed further in the paper.

Findings

The major outcome of the paper is that the COSIMA DSS contains qualities that make it suitable for handling complex assessment problems by incorporation of relevant MCDA-criteria and applications based on weights. In this way the methodology behind COSIMA sets out guidelines for dealing with the overall feasibility issues of a project appraisal by exploring whether other issues or criteria complementing the CBA can make a project change from being non-feasible to attractive, see Figure 4.1. Furthermore, the approach with its new features may be perceived as being easier accessible by the decision-makers than more complex types of MCDA.

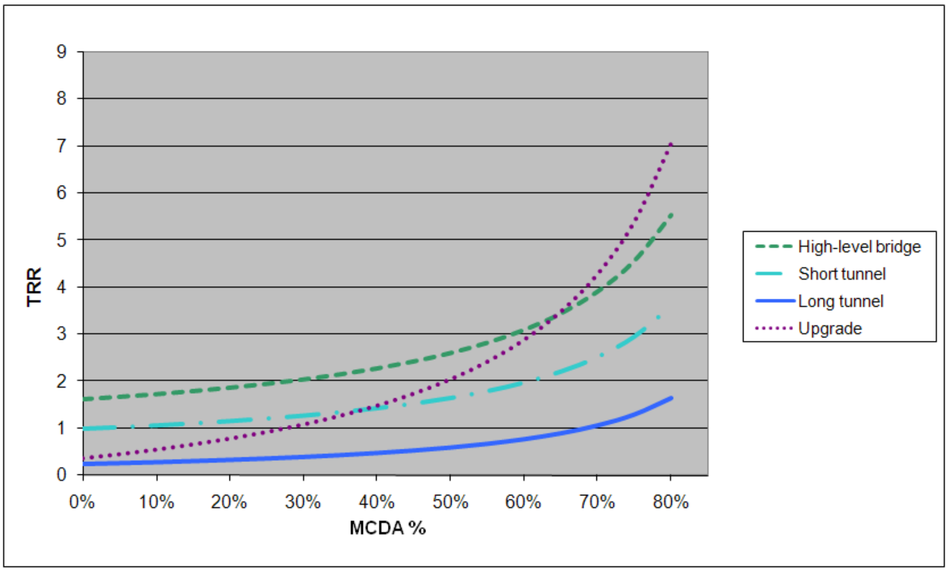


Figure 4.1. Attractiveness of the four alternatives examined as a function of the MCDA's total influence on the composite appraisal (the MCDA %)

The COSIMA DSS differs from previous attempts on doing composite appraisals in the transport sector in several ways. First of all the COSIMA DSS seeks to "translate" the MCDA results into the same "language" as the CBA results making it possible to produce a total rate of return (TRR), whereas most recent methodologies incorporate the CBA in the MCDA. Obviously, the TRR outcome from the composite expression has no economic argument even though expressed similar to the benefit-cost rate. Instead the TRR describes the attractiveness of the alternative seen from both the CBA and MCDA. Thus, the innovative advantage of using the COSIMA approach is that the CBA results are maintained throughout the analysis. Moreover, COSIMA has the advantage that expressing the outcome on a graph as depicted in Figure 4.1 makes it possible to review the results' sensitivity with regard to the weights assigned to the CBA and MCDA respectively. Overall, it is concluded that COSIMA contributes in a new way to make decisions more informed.

4.2 Paper 2

Examination of decision support systems for composite CBA and MCDA assessments of transport infrastructure projects

Author(s): Barfod, M.B., Vestergaard, A.V. and Leleur, S.

Published in Y. Shi et al. (eds). *New State of MCDM in the 21st Century – Selected papers of the 20th International Conference on Multiple Criteria Decision Making 2009*. Lecture Notes in Economics and Mathematical Systems, Volume 648, pp. 167-176, Springer 2011.

Description

The paper examines two decision support systems (DSS), REMBRANDT and COSIMA, with the purpose of identifying the most appropriate DSS for transport infrastructure assessments including both CBA and MCDA. The first DSS examined, which is widely used and based on an acknowledged methodology, comprises the REMBRANDT technique using pair wise comparisons for rating of the alternatives and determination of the criteria weights. The results of the CBA are in this system compared in a pair wise way and included as an additional criterion in the MCDA. Hence, the result of the system is a relative weight-score for each alternative reflecting its performance in the composite appraisal.

The second DSS examined – the COSIMA approach – provides a framework for adding value functions determined in a MCDA to impacts monetarily assessed in a CBA. The DSS comprises the REMBRANDT technique using pair wise comparisons for rating of the alternatives and swing weights for the determination of criteria weights. However, the COSIMA system does not convert the CBA into an additional MCDA criterion. Instead the value functions computed in the MCDA are added to the CBA results using a trade-off indicator assigning shadow prices to the MCDA criteria. Subsequently, the resulting total value is divided by the investment costs. Hence, the result is a total rate for each alternative reflecting its attractiveness in the appraisal as a function of the weight-set between the CBA and MCDA.

The input for the two DSS examined was generated using a case study concerning four corridor alternatives for a new high-speed railway line in Sweden, namely a “red” corridor (R), a “blue” corridor with a short tunnel (BS), a “blue” corridor with a long tunnel (BL), and a “green” corridor (G). For this purpose a decision conference was set up where various stakeholders and decision-makers under the guidance of a facilitator were producing input to the DSS in form of their preferences.

Findings

The major outcome of the paper is that the two DSS have been applied to the same case study revealing the same results. In Table 4.1 the rankings of the alternatives at different weight-sets are depicted for both DSS. This clearly shows that the two DSS also provide the same results on second and third level in the rankings.

Table 4.1. Rankings of the alternatives at different weight-sets

CBA weight	**0.85**		**0.70**		**0.55**	
	REMBRANDT	COSIMA	REMBRANDT	COSIMA	REMBRANDT	COSIMA
R	2	2	3	3	3	3
BS	1	1	1	1	1	1
BL	3	3	2	2	2	2
G	4	4	4	4	4	4

The difference on the results of the two DSS, however, was found to consist in the way they are expressed. The REMBRANDT DSS provides the decision-makers with weight-scores expressing the alternatives relative performance against each other. The COSIMA DSS on the other hand provides the decision-makers with a somewhat more informed result. The total rate (TRR) from COSIMA features both the CBA result and the MCDA result expressed in one single rate.

The REMBRANDT system is a theoretically well founded system which has been applied to various decision problems, and on which other systems can be measured. Given that the COSIMA DSS provides the same results (in this case) as REMBRANDT the DSS seems most appropriate for use within transport infrastructure planning as the results provide the decision-makers with two-way information containing both an economic argument and a strategic argument. This makes the COSIMA results more useful especially when the results need to be transparent and justifiable to the public.

4.3 Paper 3

Customised DSS and decision conferences

Author(s): Barfod, M.B. and Leleur, S.

Published in *the Proceedings of the 13th Euro Working Group on Transportation Meeting.* University of Padova 2009.

Description

The paper presents and exemplifies a combination of techniques for deriving and modelling decision-maker and/or stakeholder preferences using a decision conference process. The applied techniques are used for the development of customised decision support systems (C-DSS) which can be used for appraisals of large transport infrastructure projects. The paper exemplifies how the process at a decision conference can be effectively supported by a DSS customised using appropriate techniques for the specific task in hand. In this respect a conventional cost-benefit analysis (CBA) is combined with a multi-criteria decision analysis (MCDA) featuring the REMBRANDT and the swing weight techniques. The approach is presented based on a case study, which concerns the interaction between stakeholders and decision-makers at a decision conference which was set up for the appraisal of proposals for the alignment of a high-speed railway line in Sweden.

The concept of a decision conference is introduced into transport infrastructure planning as the decisions to be made in this context often are of a very complex character. This entails from the stakeholders and authorities who seem to have great leverage in the debate concerning these types of projects and hence also in the final decision. Thus a need has arisen for a structured decision process which can take all aspects into account and at the same time be transparent both to the participants and the public. The paper provides a proposal for how such a process can be prepared and designed with regard to both the decision-maker interaction and the techniques used in the underlying C-DSS.

For illustration of the process a case study concerning four corridor alternatives for a new high-speed railway line in Sweden is presented. The DSS introduced is customised to the specific assessment task using techniques that reflect the current needs and composition of the decision-makers and/or stakeholders participating in the decision process.

Findings

Customising a DSS to fit the specific assessment task in hand places high demands on the decision analysts to indentify and use the most appropriate techniques. In the case study the decision conference was attended by persons with expert knowledge about the decision problem. Consequently, it was possible to make use of a demanding technique like swing weights. The pair wise comparison technique is not a demanding technique to apply and does not require expert knowledge. However, it is a very appropriate technique when dealing with decision problems in a "local" system, where the task is to identify the best alternative for a given project – not to compare projects of different types or with geographical different locations. Setting up the C-DSS is a crucial task for the decision analysts as techniques that are appropriate for one decision problem might be inappropriate for another problem; thereby the participants' input will not be treated in a suitable manner and the assessment will in worst case provide misrepresenting results.

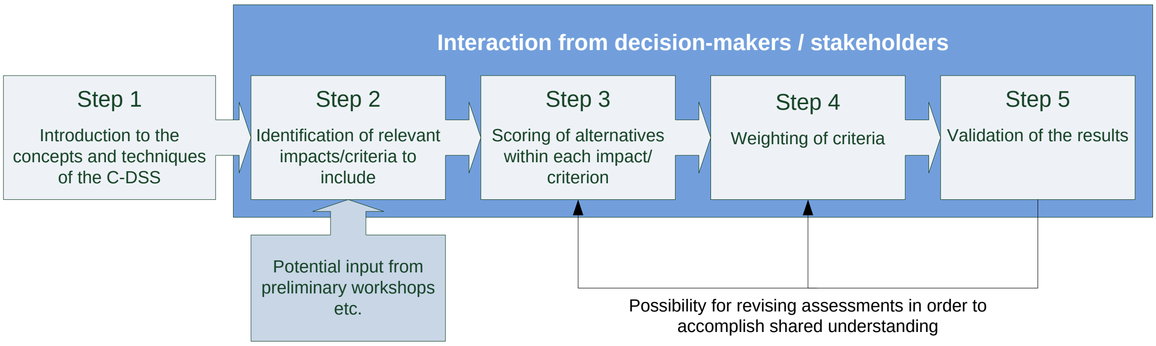


Figure 4.2. The process at the decision conference comprising five steps

The decision conference has been proposed to be set up as a five-step process leading the participants through the decision process in an easily accessible and transparent manner, see Figure 4.2. It is exemplified through the case study that the proposed C-DSS in combination with the process of a decision conference is an effective decision aid when complex decisions regarding transport infrastructure projects have to be made. This type of decision problem involves a high number of different stakeholders and decision-makers and a structured process capturing all aspects of the issue is therefore needed. In this respect the proposed methodology provides a customised process which seeks to give everybody an opportunity to express their preferences and influence the outcome. Hence, the C-DSS makes it possible for the decision-makers to make a more informed decision than would be the case if a decision conference was not carried out.

4.4 Paper 4

An MCDA approach for the selection of bike projects based on structuring and appraising activities

Author(s): Barfod, M.B.

Published in *European Journal of Operational Research* 218, Issue 3, pp. 810-818, Elsevier 2012.

Description

The paper presents an approach for the structuring and appraising of large and complex decision problems using multi-criteria decision analysis (MCDA). More specifically, the paper examines the three-step structuring process for decision analysis proposed by von Winterfeldt and Edwards (2007): 1) identifying the problem; 2) selecting an appropriate analytic approach; and 3) developing a detailed analytic structure. For illustration of the approach a case study dealing with the assessment task of structuring and prioritising initiatives and projects in a public pool with limited funds is examined throughout the paper. The process is embedded in a decision support system (DSS) making use of the REMBRANDT technique for pair wise comparisons to determine project rankings. A process for limiting the number of pair wise comparisons to be made in the process is in this connection presented.

The paper presents the efforts of the work with structuring and appraising the public Danish pool for more bike traffic – the so-called CPP case – which was conducted in late 2009 as consultancy for the Danish Road Directorate. The aim of the pool was to move users from car traffic, but also public transportation, to bikes. The Bike pool was open for applications of widely varying characters, and in principle it was possible for everybody to apply for subsidies from the pool. As a result of this the submitted applications were described in very different levels of details ranging from fully impact calculated projects to superficial map drawings of where to place a bike path. In total 133 project applications were submitted from municipalities, regions, organisations, companies and research institutions. Hence, there was a need for an appraisal of which projects should be given subsidies from the pool, as it was impossible to give subsidies to all the projects. The technical evaluation task was henceforth to design and apply a series of principles and methods which were capable of handling this large quantity of very different projects in an appropriate and optimal way in order for the total means of the pool to be allocated to those projects and initiatives that contributed the most to the overall objective.

Findings

The paper describes the structuring and appraising activities associated with a major decision analysis of projects to promote biking activities in Denmark – the CPP case – and shows decision analysis using MCDA as a useful approach for structuring and appraising large and complex decision problems. The hierarchical structure of the specific case study is depicted in Figure 4.3.

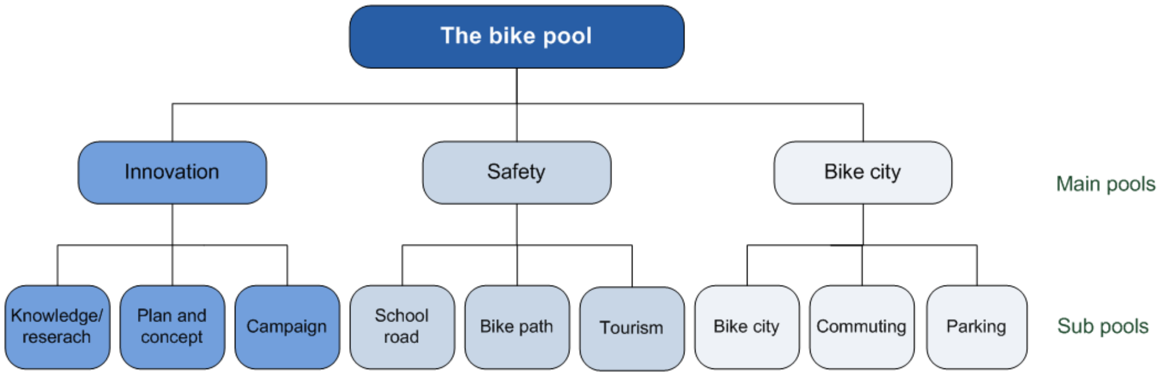


Figure 4.3. Structure of the decision problem

Some specific findings of the paper related to the structure of the decision problem and the proposed DSS are:

1. The structuring task should be conducted in close dialogue between the analysts, the decision-makers and the stakeholders.

2. Focus should be on solving the problem, not forcing a particular analytic structure onto the problem.

3. A good structure emerges when social and technical facilitation skills are combined.

4. The CPP-DSS featuring the REMBRANDT technique has shown to be a useful tool when dealing with large complex problems that are in need of a clear structure in order to be solved.

5. It is seen as a major feature of the CPP-DSS that the various inputs needed from the decision-makers can help generate important discussions in the group.

6. Future developments of the CPP-DSS should work towards inclusion of CBA as the necessary socio-economic foundation become available from research on this topic.

4.5 Paper 5

Scaling transformations in the REMBRANDT technique: a sensitivity examination of the progression factors

Author(s): Barfod, M.B. and Leleur, S.

Presented at the 21st International Conference on Multiple Criteria Decision Making, Jyväskylä, Finland, June 2011

Submitted for publication in *International Journal of Information Technology and Decision Making,* November 2011 (currently under review)

Description

The paper examines a decision support system (DSS) for the appraisal of complex transport infrastructure decision problems using multi-criteria decision analysis (MCDA). The DSS makes use of a structured hierarchical approach featuring the multiplicative AHP also known as the REMBRANDT technique. The technique is a further development of the original AHP and it proposes to overcome three issues regarding the theory behind AHP namely by using direct rating on a geometric scale, the geometric mean method, and aggregation of scores by the product of alternative relative scores weighted by the power of weights obtained from analysis of the hierarchical elements above the alternatives. The aim of the paper is mainly to address the first issue regarding the direct rating on a geometric scale.

More specifically, the paper addresses the influence of the progression factors used when transforming the decision-makers' verbal responses from the semantic to the geometric scale. The REMBRANDT technique uses the progression factor 2 for calculating scores of alternatives and $\sqrt{2}$ for calculation of criteria weights. Tests are conducted on the magnitude of the progression factors in order to examine the sensitivity towards the final outcome of an analysis.

For illustration of the DSS and the sensitivity calculations a case study dealing with the appraisal task of a large transport infrastructure project is presented. The scope of the case study is to identify the most attractive alternative for a new bridge or tunnel connection between the cities of Elsinore (Helsingor) in Denmark and Helsingborg in Sweden, which is supposed to take over both person and freight transport from the existing ferries and relieve the existing fixed link between Copenhagen in Denmark and Malmo in Sweden.

Findings

Both the original AHP and the REMBRANDT techniques can be considered as effective DSSs for group decision making, and the final scores calculated by the two versions do not strongly diverge.

In the ease of use the two versions are very similar as they need the same type of input and provide the same type of output. The original AHP has one scale only and ignores scale dependence, whereas the REMBRANDT technique, based on a one parametric class of geometric scales, yields a scale-independent rank order of the final scores and avoids rank reversal in some notorious cases where this phenomenon is not expected to occur.

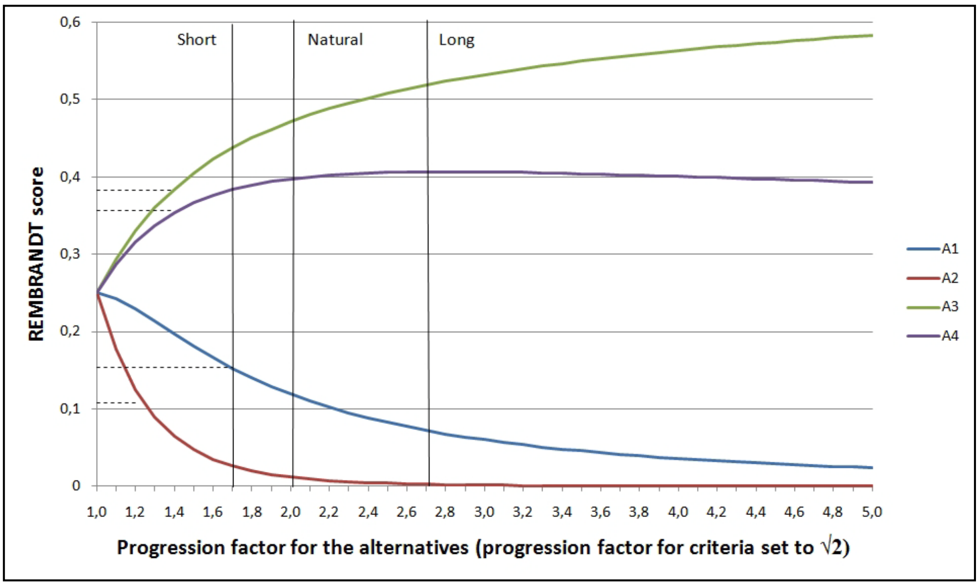


Figure 4.4. REMBRANDT scores at varying values of the progression factor for the alternatives. The denominations "short", "natural", and "long" indicate the scales proposed for sensitivity analysis, and the horizontal dashed lines indicate the scores calculated using the original AHP.

Based on the main case study and further tests it can be recommended to conduct sensitivity analysis applying different progression factors on both the alternatives scores level and the criteria weights level. The rank order of the alternatives does not depend on the scale parameter γ when this is changed for the alternatives scores level. However, it can be concluded that the scale parameter should not exceed a long scale (a progression factor on 2.7) by much as the span between the scores becomes inappropriately large. As opposed to this the rank order of alternatives are very sensitive towards changes in the progression factor on the criteria weights level. Therefore it can be recommended that a progression factor on 1.7 (a short scale) can be applied at the alternatives scores level while the criteria weights level should continue to make use of a progression factor on√2 if it is desirable to arrive at results closer in line with the original AHP, see Figure 4.4.

4.6 Impact of the papers

As all five papers deal with decision support within transport planning it is appropriate to discuss how they differ and what their combined contribution is. The papers all serve the purpose of introducing MCDA in the appraisals of transport infrastructure projects. Concerning the time frame of the study three phases have been scrutinised in which the decision support framework has undergone substantial changes.

Paper 1, even though published in 2011, was originally written in early 2008. Herein the focus was to present and further develop the COSIMA methodology of combining CBA and MCDA. Several techniques were introduced and implemented within the assessment framework. In paper 2 the COSIMA methodology was further explored and a validation of the methodology was made by comparing its results with the results of the well-founded REMBRANDT technique applied on the same case study. Paper 3 and 4 were focussed on the decision process and how to derive preferences from the decision-makers. In this respect the concept of decision conferences was explored and a framework for the decision process was set up comprising five steps for the decision-makers to explore. The process was moreover taken a step further by using a decision analytic approach to structure a complex decision problem in order to make it possible for the ratifying group to conduct an appropriate assessment at the decision conference. Paper 5 made a thorough exploration of the REMBRANDT technique and its differences with the original AHP technique. In particular the transformation when going from the verbal to the numerical scale was addressed and a modification of the progression factors was proposed.

The main focus of this Ph.D. study has first been to develop a process and framework for providing valid, flexible and effective decision support in situations where complex decision problems concerning transport infrastructure projects are to be assessed. In this respect a major concern has been to bring informed and transparent decision support to the decision-makers in terms of the output of the analysis. Second, emphasis has been on identifying the most appropriate assessment techniques for the decision process in order to make the judgments to be made as simple and transparent as possible for the decision-makers.

Table 4.2 depicts the way each research outcome has been handled through each paper. The table forms the basis for the concluding remarks of this Ph.D. thesis.

Table 4.2. Research outcome of the five papers

Main focus	**Paper 1**	**Paper 2**	**Paper 3**	**Paper 4**	**Paper 5**
1. The composite model for assessment (COSIMA) is an effective decision support system (DSS) for complex planning problems involving both monetary impacts and non-monetary criteria.	x	x			
2. Direct rating using pair wise comparisons is found to be an appropriate MCDA approach for computing scores for alternatives while rank based approaches are appropriate for eliciting criteria weights from the decision-makers' preferences.		x	x	x	
3. Decision analysis and decision conferences using MCDA are useful approaches for structuring and appraising large and complex decision problems with participation of relevant stakeholders and decision-makers.			x	x	
4. The REMBRANDT technique with a modified progression factor can be recommended for practical use instead of the original AHP to derive decision-maker preferences.				x	x

5 Findings relating to assessment techniques and examination process

This chapter presents the findings of this Ph.D. thesis relating to both the assessment techniques and the examination process for the use in decision making for transport infrastructure projects. The thesis has focussed on two main issues: if it is possible to propose an examination process that can be used in situations where complex decision problems need to be addressed by technicians as well as decision-makers and citizens, and how the methodologies and techniques made use of within the examination process can be optimised to meet the specific decision task in hand.

The findings should not be seen as the only or definite answer to how to obtain an optimal decision making process which can lead to a possible decision. Instead they are first steps on the way towards achieving this. Even though, the findings should be seen as a help to transport planners dealing with complex decision problems involving various decision-makers and stakeholders. Further development is needed in order to make them comprehensive enough to cover the often changing planning conditions. However, it will most likely never be possible to set up a process and select assessment techniques that will provide suitable solutions to all types of decision situations and problems.

5.1 The assessment techniques

It is necessary in the beginning of a decision process to select a specific analytical approach for the development of the appraisal framework. This model building can be regarded as a dynamic process which interacts with the process of the appraisal. The nature of the analytic approach which is selected will differ according to the nature of the assessment and the definition of the alternatives. Moreover, the composition of the ratifying group using the approach should also be considered: are we dealing with professionals/experts or persons with only a superficial knowledge about the decision problem? The task of selecting an assessment technique might very well lead to the realisation that one technique is not sufficient to meet the requirements of the decision problem. For this reason a mix of techniques might be a useful solution. For transport infrastructure projects in particular it is a requirement in Denmark that a conventional cost-benefit analysis (CBA) is conducted which then can be complemented with strategic impact assessments in a multi-criteria decision analysis (MCDA). This section is limited to findings concerning when the examined methods belonging to the normative approach (American school) within MCDA should be used.

It is most likely that different decision situations can occur containing some of the following characteristics:

- The alternatives to be assessed can either be well-defined and easy to measure with regard to potential impacts or they can be poorly defined making them very difficult to assess.
- The criteria to be weighted can either be based on well-defined measureable attributes and easy to weight, or the attributes can be non-measureable making it difficult to interpret the weights.
- The decision-makers or the ratifying group can either consist of professionals which are experts within their area and have experience in the type of judgments to be made, or they can be persons with only a superficial level of knowledge about the issue in hand.

The different decision situations sketched above set varying requirements to the techniques to be used. The following techniques have been treated in this Ph.D. thesis and their recommendations are as follows:

- **SMART** (Simple Multi-Attribute Rating Technique) is based on the additive value function model and assigns direct scores to alternatives and direct weights to criteria. The technique demands a high level of knowledge about the alternatives/criteria to be assessed and should for that reason only be used when measureable attributes can be identified. Moreover, the technique should only be used by experts or professionals which are experienced users of the technique.
- **Swing weights** is usually considered to be the theoretical most correct method for deriving criteria weights, but it is most likely also the most difficult one to use in practice. The technique presupposes that the decision-makers consider the swing from worst to best within each criterion and rank the criteria based on which swing gives the highest increase in overall value. Afterwards the swings within each of the criteria are assigned with a numerical value reflecting its importance compared to the swing within the most important criterion. In practice the technique is difficult to explain to non-professionals and should for that reason only be used with care.
- **SMARTER** (Simple Multi-Attribute Rating Technique Exploiting Ranks) is, as the name implies, a further development of SMART. The technique is very simple in the sense that it only demands the decision-makers to rank the criteria in order of importance after which predetermined surrogate weights are assigned to the criteria, e.g. ROD weights. The method presupposes no measurable attributes and is easy accessible and very simple to use for decision-makers which are non-professionals. In practice the technique has, moreover, shown to meet the decision-makers' requests.
- **AHP** (Analytic Hierarchy Process) offers an alternative to SMART that is based on pair wise comparisons of alternatives and criteria to obtain scores and weights. A nine

point intensity scale of importance is in this context used to express the decision-makers' preference for one object over another. AHP is very simple to use compared with SMART as the problem is decomposed into simple judgments requiring no measurable attributes. The technique is very useful in situations where the alternatives are weakly described and where it is difficult to assign weights to the criteria. Moreover, the technique has proven its worth in group decision making situations where scores and weights are obtained through discussions.

- **REMBRANDT** (Ratio Estimations in Magnitudes or deci-Bells to Rate Alternatives which are Non-DominaTed) is a further development of AHP based on the multiplicative value function model. REMBRANDT offers a more theoretically correct approach than AHP, but in practice the two methods demand the same type of input and generate the same type of output. Thus REMBRANDT is useful in the same decision situations as AHP. However, it is recommended to use REMBRANDT instead of AHP as the only differences on the techniques consist in the theoretical foundation; the input to the techniques are derived identically and the output provided expresses the same results.

- **COSIMA** (COmpoSIte Model for Assessment) is a methodology for combining CBA and MCDA into one single rate of attractiveness. The set-up of the model only requires that value function scores can be assigned to the alternatives and weights can be determined for the criteria. Hence, no specific MCDA technique is linked to the methodology, and it is possible to apply all the mentioned techniques as well as a combination of them dependent of what seems relevant for the decision problem in hand.

Based on the above it is clear that different techniques should be used dependent on the persons to apply them in the decision process. Two main modes are in this respect relevant: a basic-user mode consisting of non-professionals, and an expert-user mode consisting of professional and experienced users of the techniques. Table 5.1 depicts the techniques that are recommended for use in the two modes.

Table 5.1. Recommended techniques

	Basic-user mode	Expert-user mode
Criteria weights	SMARTER	Swing weights
Alternative scores	REMBRANDT	SMART / REMBRANDT

Basic users should make use of the simplest set-up of techniques as possible in the decision process in order to avoid misunderstandings and misinterpretations. Thus it is recommended that the SMARTER technique is used for assigning weights to the criteria, as this only requires the users to rank the criteria in order of importance. For assigning scores to the alternatives it is recommended that the basic users should make use of the REMBRANDT technique, which requires them to consider simple pair wise comparisons according to a semantic scale. An alternative could be to use AHP if, for example, a person in the ratifying group has used the technique before and feel confident with the results of this.

Expert users must be considered to be capable of perceiving more demanding methods than the basic users as they often are professionals with much experience in assessment tasks. For this reason it is recommended that the swing weight technique is applied to determine weights for the criteria, as the technique makes it possible to determine the weights with relatively high accuracy. For the scoring of alternatives the SMART technique is recommended for use if the attributes are measureable, if not the REMBRANDT technique should be applied as in the case with the basic users.

5.2 The examination process

The examination process should always be designed to accommodate the actual infrastructure project to be appraised. Minor transport infrastructure projects may be examined using a conventional CBA whereas large transport infrastructure projects often will necessitate that strategic impacts are included and that alternative developments are considered. Throughout the thesis a methodological framework is formulated which, besides its practical purpose, also aims at presenting a wider theoretical background for the modelling of the decision-makers' preferences.

Some examination principles are formulated throughout the five papers which combined proposes an examination process concerning large transport infrastructure assessments in general. Figure 5.1 depicts this process.

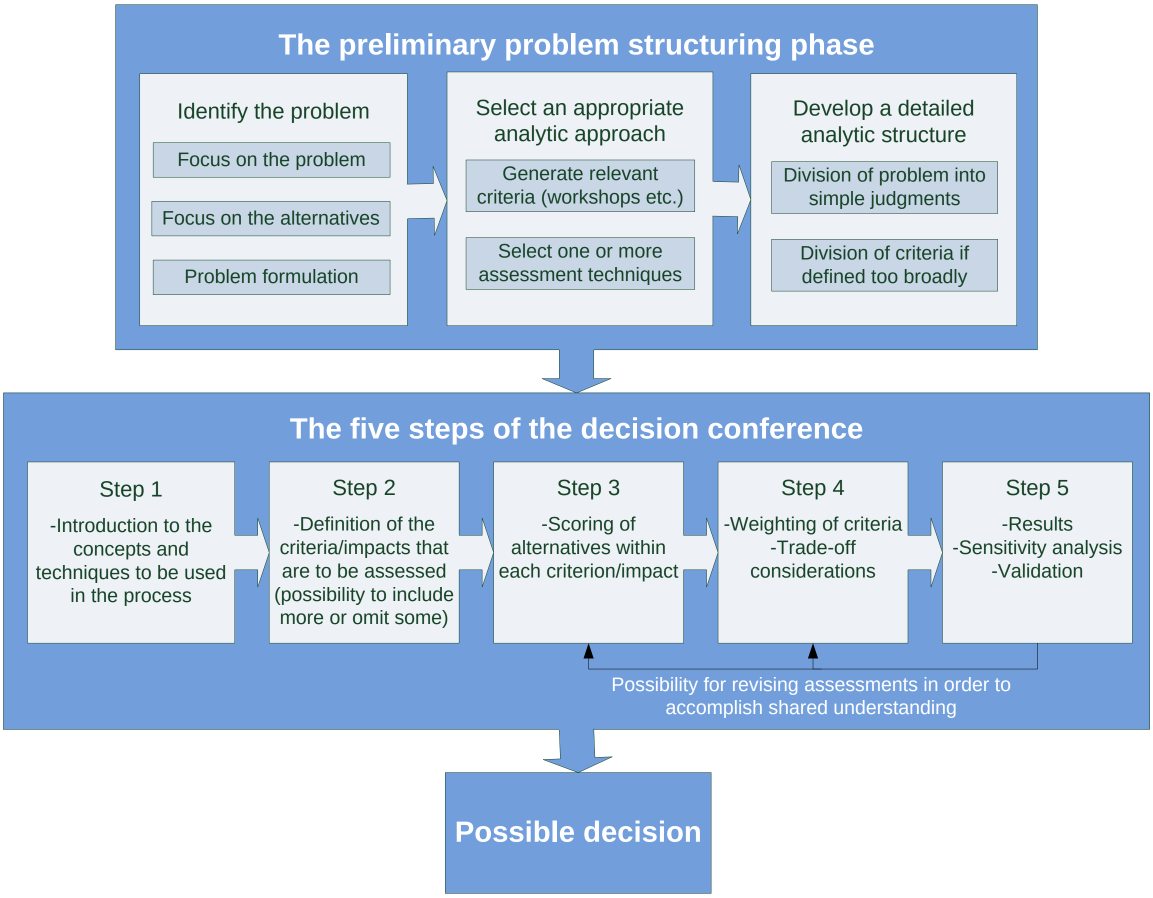


Figure 5.1. Proposed examination process

The process can be divided into two main phases: the preliminary problem structuring phase and the interaction phase (the decision conference). The preliminary phase takes its point of departure in problem structuring methods where the problem in the first stage is identified by using techniques for focussing on the problem and on the possible alternatives, and doing a problem formulation. This part is not treated in dept in this thesis, but is described thoroughly in e.g. Jeppesen (2009).

At the next stage the emphasis of the preliminary phase moves from problem structuring to model building where a specific analytical approach needs to be chosen for the development of a framework for the appraisal. In this context the model building is regarded as a dynamic process, informed by and informing the problem structuring process, and interacting with the process of appraisal. It may involve some iteration, search for new criteria (e.g. by conducting workshops), discarding, reinstating and redefining old ones, and further extensive discussions amongst the participants in the process. Moving from a broad description of the problem, whether it is a simple clustering of ideas, a fully elaborated map, or some other representation of the issue, to a preliminary definition of a model for MCDA, requires a good understanding of the chosen approach to multi-criteria modelling. The nature of the approach which is selected will, as noted in section 5.1, differ according to the nature of the assessment task and whether the alternatives are explicitly or implicitly defined. The task of selecting an assessment technique might lead to the realisation that one technique is not sufficient to meet the requirements of the decision problem. One technique might be useful for the scoring of alternatives while another technique is useful for the weighting of criteria; this depends on the problem to be assessed. Moreover, it can be a requirement that some impacts are assessed using a CBA as is the case for transport infrastructure projects.

This leads to the third stage of the preliminary problem structuring phase where the detailed analytic structure is developed using input from the two first stages. The chosen approach is used to structure the problem including all relevant criteria and alternatives creating sub-divisions if needed. The objective is to structure and decompose the problem into simple judgments for the decision-makers to consider.

Once the preliminary structuring phase has been conducted it is possible to start the actual assessment of the problem. For this purpose the decision conference approach is recommended as it makes it possible to bring different stakeholders and decision-makers together with the purpose of obtaining some common agreement on the issue. As described in Paper 3 such a decision conference can be built up by a five-step process leading the participants through the decision process. The five steps are universal and can be applied regardless of the nature of the decision problem considered; only minor adjustments should need to be made within the steps.

First, the participants of the decision conference should be introduced to the concepts and techniques to be used in the process in order for the output not to be perceived as a “black box”. Second, the criteria to be assessed should be defined in a high level of detail to avoid any misunderstandings and double counting with other criteria. In this step it will be possible to include additional criteria if there is a need for this. Third, the alternatives are scored within each criterion using the assessment technique chosen in the preliminary phase. The criteria are in the fourth step weighted and if the MCDA has to be combined with a CBA, trade-off considerations are made. Finally, results are obtained and the process is validated. If the participants are not convinced about the final results it is possible to go back in the process and revise the assessments made. However, it must be expected that it is possible to arrive at a common agreement and a possible decision after some iterations.

5.3 Discussion of findings

The main purpose of this Ph.D. study has been to develop a process and framework for providing valid, flexible and effective decision support in situations where complex decision problems concerning transport infrastructure projects are to be assessed. In this respect focus has been on proposing a process that is valid in the sense that it meets the need of the users and fulfil its purpose effectively. The proposed framework is intended to examine real decision situations.

CBA is not a stand-alone activity; it is part of a larger effort to appraise and evaluate projects as stated by the World Bank (2010). Most of the information collected for e.g. an environmental impact assessment (EIA) is useful under any approach to deliberation. The problems arise only in the final steps of aggregating everything into a single bottom-line number: monetising non-monetary benefits, discounting future outcomes, and estimating the values of important uncertainties all have the effect of distorting and concealing the underlying data. An infrastructure project proposal has monetary costs, in the present and perhaps the near future, and a mix of monetary and non-monetary benefits extending somewhat further into the future. Knowing as much as possible about the costs and – separately – about the benefits, with each expressed in its natural units, is sure to lead to better decisions.

In CBA money measures impacts in a single dimension. All impacts that have meaningful unit prices can be unambiguously compared to each other; the market leaves no doubt about which are worth the most, and which the least. In contrast, human lives and the environment involve many dimensions of incompatible measurements. One shortcoming of CBA is that it depends on a reduction of these multiple dimensions of value to the single metric of money. MCDA on the other hand accepts and builds upon a multi-dimensional set of objectives. Rather than a single score or market value, MCDA evaluates projects and alternatives by multiple standards – often four to eight criteria, although the number can vary. If a project looks good on some but not all criteria, MCDA can report this finding in easily understandable terms, whereas CBA would tend to hide the contrasting patterns of results from view. As noted in Leleur (forthcoming) based on the short-comings of market pricing, CBA can be argued to be an economics approach, while MCDA is more an engineering approach.

The strength of MCDA is its transparency in reporting complex evaluations, where the result is not entirely black or white. It is particular well-suited to situations where different stakeholders emphasise different objectives, offering a formula for calculating and presenting the issues. The weakness of MCDA is, however, its reliance on decision-maker judgments about the criteria. Both the choice of criteria to be included and the relative weight given to each criterion will influence the final result. There will often be too many possible criteria to include them all; choices about inclusion and exclusion must inevitably be made, and the question of how important the criteria are in relation to each other needs to be addressed.

This thesis has in particular made use of the pair wise comparison techniques: AHP and REMBRANDT. Despite these techniques' many positive aspects, such as simplicity and transparency, a major downside is the number of comparisons that the respondent group has to make. The more alternatives in an assessment the more inappropriate the pair wise comparison technique becomes. If there are too many comparisons to be made the respondent group tends to get tired and thereby make comparisons of a lower quality; the comparisons can then tend to be taken as more or less an average of the groups' viewpoints. This will influence the rest of the appraisal and generate poor results. In order for this not to be an issue another assessment technique such as SMART should be chosen if the number of comparisons exceeds what seems reasonable to manage within the given time frame of the assessment task.

Furthermore, it is most likely that the weightings of the criteria will diverge depending of the decision-maker or stakeholder that makes the judgments. Thus, this part of the assessment can be seen as more subjective than the part concerning the alternatives where the issue can be broken down into simple objective judgments within each criterion. Therefore, it can be beneficial to evaluate different stakeholders' preferences in order to obtain a broader perspective for the final decision making.

Despite these limitations, MCDA is still the right choice for many situations where clear, conflicting objectives must be discussed and reconciled, and where stakeholders are stating different preferences. It is important to be aware of the limitations of the method when using it and to recognise the influence of the factual decisions behind the framing of the issue.

The existing appraisal framework in Denmark (DMT, 2003) does not attempt to incorporate the strategic issues (the MCDA-criteria) of a decision problem into appraisals of transport infrastructure projects. Other frameworks, such as the EUNET (2001) framework, incorporate the CBA results as a criterion in the MCDA and the result of the framework is expressed in form of a relative rate. Using the COSIMA approach the decision-makers are provided with a result that contains a level of information which comprises both the CBA and MCDA expressed in a more easily accessible way. Generally, decision-makers are used to make decisions on the basis of a B/C-rate and are hence comfortable with this type of expression. The basic feature in the COSIMA approach is that the MCDA part is converted to the same scale as the CBA part providing the decision-makers with an indication of the value of the strategic issues based on their own preferences expressed as a TRR. The TRR result will most likely vary based on who is stating the preferences; however, by assuring diversity in the assessment group the result becomes valid to a wide audience.

An important issue to address in this context is again the importance of the criteria weights. The weights are directly linked to the shadow prices assigned to the criteria in the COSIMA approach and changes in the weights will thus have a large influence on the final outcome of the analysis. The criteria weights can, as mentioned, be seen as the most subjective part of the MCDA assessment and will differ depending on who is setting them. For this reason sensitivity analysis should be conducted testing different weight sets in order to see how changes will affect the investment decision to be taken.

The use of local value function scales within the COSIMA system also raises some issues. First, projects assessed using this scale cannot easily be compared with other projects as the scale presupposes a closed system. This way the endpoints within the assessment define the scale, in contradiction to a global scale which considers the extreme endpoints in defining the scale. If the projects are not assessed using the same scale it is obvious that they cannot be compared to each other. Second, the segregation between the alternatives within some criteria can in some cases almost be negligible. A means to overcome this problem is to perform a check of all the criteria in the assessment using the swing weight method. If this check reveals some criteria where the swing from the best to the worst performing alternative almost does not exist (a lower boundary can be set by the decision-makers), the criteria should be omitted from the analysis as they do not contribute to the segregation between the alternatives and therefore are without significance for the appraisal task in hand.

The decision conference approach has been proposed as a process for structuring the decision process, where all relevant stakeholders can participate and influence the results. An important aspect in this context is the documentation of the assessments and choices made along the way. This is especially the case if the outcome of the decision conference has to be justified to third parties (e.g. the public) where thorough and transparent argumentation is needed. In this respect an assessment protocol can be very useful to record the rationale of the statements made during the decision support process. If inconsistencies occur in the assessments the protocol can be helpful to clear out misunderstandings or errors, and corrections can be made effectively using the recorded rationale. Moreover, the protocol can include notes about possible disagreements in the group with regard to the assessments and how these were dealt with. A proper documentation of the decision conference can be very helpful both with regard to the conference itself but also when the outcome has to be justified. The protocol should therefore be seen as an integrated part of the decision conference approach.

Decisions made in consensus at a decision conference can be expected to have a fairly higher probability for being implemented than results from a complex decision analysis that only involves one decision-maker who later has to justify his decision for other people (e.g. in a organisation or to the public). Moreover, decisions made by such groups have better terms for working in practice as they have the group’s commitment. However, the question whether decisions made in consensus at a decision conference using a customised DSS are more or less valid than assessments and solutions made without these aids has to be answered. According to Phillips (2007) this is not necessarily the case; however, it is evident that a decision conference provides some advantages regarding: better communication between groups, a common understanding of strategic objectives and hence common commitment towards the objective, improved teamwork, better knowledge and relation to various uncertainties, and finally and foremost decisions that can be defended.

As mentioned in Section 3.6 the concept of decision conferencing was originally developed in the late 1970s in USA for the use in private companies when addressing complex investment decisions. Later on the concept has then been revised and found applicable to support decisions concerning public investments as well. Decision conferences have provided a framework that could be implemented more in transport planning processes, as the decision conference outline enables a structured debate between stakeholders and decision-makers. This is an important task which is often requested in planning situations. However, as the goal of the debate is not necessarily to obtain consensus, but instead to accomplish a shared understanding and an acceptable and liveable solution for all participants, the name of the decision conference seems rather inappropriate as no decision needs to be made. In future applications it could therefore be considered to change the name of the concept from decision conference to planning workshop to emphasise the characteristics of the concept as outlined above.

6 Conclusions and perspectives

This Ph.D. study has its focus on optimising transport decision making using customised decision models and decision conferences. Optimising in this context relates both to process and to methodologies/techniques, see below. A major concern has been how conventional cost-benefit analysis (CBA) can be supplemented with effects of a strategic character, while at the same time maintaining its purpose of providing straight-forward and transparent decision support.

This Ph.D. study has been conducted both by undertaking a theoretical literature study and by practical applications using case studies spanning over different transport modes. It appears from the five papers included in this thesis that the approach to transport decision making has developed over the whole study period from early 2008 to late 2011. The treatment of optimising transport decision making was set out as two major research questions:

I. Is it possible to propose an examination process that can be used in situations where complex decision problems need to be addressed by technicians as well as decision-makers and citizens?

II. How can the methodologies and techniques made use of within the examination process be optimised to meet the specific decision task in hand?

On the basis of the discussion of the papers in Chapter 4 and the findings in Chapter 5 the following four main conclusions can be made relating to the research questions.

The COSIMA approach has been presented as a methodology that makes it possible to combine CBA and multi-criteria decision analysis (MCDA) into one comprehensive assessment. COSIMA is simple in its design and application compared to earlier attempts to composite analyses, as the methodology basically just "adds to" and does not hide or change the information given by the CBA. Furthermore, it contains qualities that make it suitable for handling complex assessment problems by incorporation of relevant MCDA-criteria and applications based on weights. The total rate of return (TRR) outcome from the composite expression, however, has no economic argument even though expressed similar to the B/C-rate. Instead the TRR describes the attractiveness of the alternative seen from both the CBA and MCDA. Thus, the advantage of using the COSIMA approach is that the CBA results are maintained throughout the analysis. The first major outcome is the validation of the COSIMA approach.

1. The composite model for assessment (COSIMA) is an effective decision support system (DSS) for complex planning problems involving both monetary impacts and non-monetary criteria.

A major concern has throughout the study been to identify appropriate assessment techniques for different decision situations. Several MCDA techniques has been examined and two types of approaches stand out, namely the pair wise comparisons featuring AHP and REMBRANDT, and the rank based approaches SMARTER and swing weights. Thus, the general conclusion and second major outcome concerns the choice of techniques.

2. Direct rating using pair wise comparisons is found to be an appropriate MCDA approach for computing scores for alternatives while rank based approaches are appropriate for eliciting criteria weights from the decision-makers' preferences.

The importance of doing a thorough structuring of the decision problem before applying assessment techniques has furthermore been emphasised. Especially, the structuring and appraising activities associated with a major decision analysis of projects to promote biking activities in Denmark – the CPP problem described in Paper 4 – shows that decision analysis using MCDA is a very useful approach for the structuring of complex decision problems. For the practical assessments after the structuring phases the decision conference approach has been proposed. This should be set up as a five-step process leading the participants through the decision process in an easily accessible and transparent manner. Through the case studies in Paper 1 – 4 it is exemplified that a customised DSS in combination with the process of a decision conference is an effective decision aid when complex decisions regarding transport infrastructure projects have to be made. This type of decision problem involves a high number of different stakeholders and decision-makers and a structured process capturing all aspects of the issue is therefore needed. The third major outcome is consequently regarding the decision process.

3. Decision analysis and decision conferences using MCDA are useful approaches for structuring and appraising large and complex decision problems with participation of relevant stakeholders and decision-makers.

The most extensively examined assessment approach throughout the study has been the application of pair wise comparisons. Both the AHP and the REMBRANDT techniques can in this respect be considered as effective DSS for group decision making. In the ease of use the two techniques are very similar as they need the same type of input and provide the same type of output. The original AHP, however, has one scale only and ignores scale dependence. The REMBRANDT technique, based on a one parametric class of geometric scales, yields a scale-independent rank order of the final scores and avoids rank reversal in some notorious cases where this phenomenon is not expected to occur. The fourth and final major outcome concerns which pair wise comparison technique should be applied in practice.

4. The REMBRANDT technique with its better theoretical foundation can with a modified progression factor be recommended for practical use instead of the original AHP to derive decision-maker preferences.

Based on examinations of all the case studies in Paper 1 – 5 it can be recommended to conduct sensitivity analysis applying different progression factors on both the alternatives scores level and the criteria weights level when using the REMBRANDT technique. It is recommended that a progression factor on 1.7 should be applied at the alternatives scores level instead of the conventional progression factor on 2, while the criteria weights level should continue to make use of a progression factor on $\sqrt{2}$.

This Ph.D. thesis provides a broad foundation for further exploration and application of a MCDA based decision support framework. It is concluded based on the findings that MCDA ought to have a more widespread use in transport planning as several types of appraisal problems can be approached in an adequate way by making use of MCDA, where process and methodology is customised (optimised) in accordance with the actual case dealt with. A number of perspectives and future research possibilities are outlined related to both the applications of MCDA techniques and the examination process. Hence, future research within this area can include the following:

- Studying the further properties of the REMBRANDT technique seen from both a theoretical but also empirical point of view.
- The decision conference approach should be developed further with respect to the interaction between the model and the participants. New technical appliances could e.g. ease the assessment process and make it possible for even more stakeholders and civil groups to have a say about the infrastructure under consideration.
- Experts and non-experts may have different perceptions of what valid decision support is. A future framework should be able to incorporate this aspect as well.

Generally, the framework consisting of the proposed examination process and the MCDA techniques can be refined based on more practical applications and new studies of various assessment problems in transport planning. In this respect the framework set out as a result of this Ph.D. work is seen as a suitable platform.

References

Ackerman, F. and Heinzerling, L. (2002). Pricing the priceless – Cost-Benefit Analysis of environmental protection. *University of Pennsylvania Law Review* 150, pp. 1553-1584.

Alcock, P. and Powel, M. (2011). *Welfare theory and development*. Sage Publications, London, United Kingdom.

Bana e Costa, C. A. and Vansnick, J. C. (1994). MACBETH – an interactive path towards the construction of cardinal value functions. *International Transactions in Operational Research* 1, pp. 489-500, Wiley.

Banister, D. and Berechman, J. (2000). *Transport Investment and Economic Development*. UCL Press, London, United Kingdom.

Barron, F. H. and Barret, B. E. (1996). Decision Quality using Ranked Attributes Weights. *Management Science* 42 (11), pp. 1515-1523, Informs.

Barzilai, J., Cook, W. and Golany, B. (1987). Consistent weights for judgement matrices of the relative importance for alternatives. *Operations Research Letters* 6:3, pp. 131-134, Elsevier.

Beder, S. (2000). Costing the Earth: Equity, sustainable development and environmental economics. *New Zealand Journal of Environmental Law* 4, pp. 227-243, University of Auckland.

Bell, D., Raiffa, H. and Tversky, A. (1988). *Decision Making: Descriptive, Normative and Prescriptive Interactions* (Eds). Cambridge University Press, New York, United States of America.

Belton, V. and Stewart, T. J. (2002). *Multi Criteria Decision Analysis: An integrated approach*. Kluwer Academic Publishers, London, United Kingdom.

Beukers, E., Bertolini, L. and Te Brömmelstroet, M. (2012). Why Cost Benefit Analysis is perceived as a problematic tool for assessment of transport plans: A process perspective. *Transportation Research Part A* 46, pp. 68-78, Elsevier.

Chun, K-J. (1992). The effectiveness of a facilitated group decision support system (decision conferencing): a UK/US field study. London School of Economics.

Dasgupta, A. K. and Pearce, D. W. (1978). *Cost Benefit Analysis*. The Macmillan Press Ltd., London, United Kingdom.

Danish Ministry of Transport (2003). Manual concerning socio economic analysis – applied methodology in the transport sector (in Danish). Danish Ministry of Transport, Copenhagen, Denmark.

Edwards, W. and Barron, F. H. (1994). SMARTS and SMARTER: Improved Simple Methods for Multiattribute Utility Measurement. *Organizational Behavior and Human Decision Processes* 60, pp. 306-325, Elsevier.

Edwards, W., Miles, R.F. and von Winterfeldt, D. (2007). *Advances in Decision Analysis – From Foundations to Applications.* Cambridge University Press, New York, United States of America.

Engelbrecht, H-J. (2009). Natural capital, subjective well-being, and the new welfare economics of sustainability: Some evidence from cross-country regressions. *Ecological Economics* 69, pp. 380-388, Elsevier.

EUNET (2001). EUNET/SASI Final Report – Executive Summary. Report, 4th RTD Framework Programme of the European Commission.

Gissel, S. (1999). Decision aid methods in rail infrastructure planning. Ph.D. thesis, Report 4, Department of Planning, Technical University of Denmark.

Goklany, I. M. (2009). Discounting the Future. *Regulation* 32, pp. 36-40, The Sheridan Press.

Goodwin, P. and Wright, G. (2009). *Decision Analysis for Management Judgment*. Fouth Edition, Wiley, Chichester, United Kingdom.

Hobbs, B. F. and Meier, P. (2000). *Energy Decisions and the Environment – A Guide to the Use of Multicriteria Methods*. Kluwer Academic Publishers, Massachusetts, United States of America.

Hwang, C. L. and Yoon, K. (1981). Multiple Attribute Decision Making – Methods and Applications. *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag.

Hwang, C. L. and Yoon, K. (1995). Multi Attribute Decision Making – An Introduction. *Sage University Papers*, Series/Number 07-104, Sage Publications, pp 59-66.

Jeppesen, S. L. (2009). Sustainable Transport Planning – A Multi-Methodology Approach to Decision Making. Ph.D. thesis, Department of Transport, Technical University of Denmark.

Kahneman, D. and Knetsch, J. L. (1992). Valuing Public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management* 22, pp. 57-70, Academic Press.

Keeney, R. L. (1992). *Value Focussed Thinking – A Path to Creative Decisionmaking*. Harvard University Press, Massachusetts, United States of America.

Keeney, R. L. and Raiffa, H. (1993). *Decisions with Multiple Objectives*. Cambridge University Press, New York, United States of America.

Leleur, S. (2000). *Road Infrastructure Planning – A Decision-Oriented Approach*. 2nd Edition, Polyteknisk Press, Kgs. Lyngby, Denmark.

Leleur, S., Holvad, T., Salling, K. B., and Jensen, A. V. (2004). Development of the CLG-DSS Evaluation Model. Technical Report, Centre for Logistics and Freight Transport.

Leleur, S., Petersen, N.B., and Barfod, M.B. (2007). The COSIMA Approach to Transport Decision Making – Combining Cost-benefit and Multi-criteria Analysis for Comprehensive Project Appraisal. In: Kim, K-S. (Ed.). *Improving Public Investment Management for Large-Scale Government Projects: Focusing on the Feasibility Studies – Lessons and Challenges* (pp. 100-122). KDI.

Leleur, S. (forthcoming). *Complex Strategic Choices – Applying Systemic Planning for Strategic Decision Making.* Series in Decision Engineering, Springer-Verlag.

Lootsma, F.A. (1988). Numerical scaling of human judgement in pairwise-comparison methods for fuzzy multi-criteria decision analysis. In: G. Mitra (Ed.). *Mathematical Models for Decision Support* (pp. 57-88). Springer-Verlag.

Lootsma, F. A. (1992). The REMBRANDT system for multi-criteria decision analysis via pair wise comparisons or direct rating. Report 92-05, Faculteit der Technische Wiskunde en Informatica, Delft University of Technology, Netherlands.

Lootsma, F.A. (1993). Scale sensitivity in the multiplicative AHP and SMART. *Journal of Multi-Criteria Decision Analysis* 2, pp. 87-110, Wiley.

Lootsma, F. A. (1999). *Multi-Criteria Decision Analysis via Ratio and Difference Judgment*. Kluwer Academic Publishers, Dordrecht, Netherlands.

Mustajoki, J., Hämäläinen, R. P. and Marttunen, M. (2004). Participatory multicriteria decision analysis with Web-HIPRE: a case of lake regulation policy. *Environmental Modelling & Software* 19, pp. 537-547, Elsevier.

McCartt, A. and Rohrbough, J. (1989). Evaluating Group Decision Support System Effectiveness: A Performance Study of Decision Conferencing. *Decision Support Systems* 5, pp. 243-253, Elsevier.

Næss, P. (2006). Cost-benefit analysis of transportation investments – neither critical nor realistic. *Journal of Critical Realism* 5, pp. 32-60, Brill.

Olson, D. L., Fliedner, G. and Currie, K. (1995). Comparison of the REMBRANDT system with the analytic hierarchy process. *European Journal of Operational Research* 82, pp. 522-539, Elsevier.

Olson, D. L. (1996). *Decision Aids for Selection Problems*. Springer Series in Operations Research, Springer-Verlag, New York, United States of America.

Osland, O. and Strand, A. (2010). The politics and Institutions of Project Approval – a Critical-Constructive Comment on the Theory of Strategic Misrepresentation. *European Journal of Transport and Infrastructure Research* 10, pp. 77-88, Technical University of Delft.

Phillips, L.D. (1984). A theory of requisite decision models. *Acta Psychologica* 56, pp. 29-48, Elsevier.

Phillips, L.D. (2007). Decision conferencing. In: Edwards, W., Miles, R.F. and von Winterfeldt, D. (Eds.). *Advances in Decision Analysis – From Foundations to Applications* (pp. 375-399). Cambridge University Press, New York, United States of America.

Roberts, R. and Goodwin, P. (2002). Weight Approximations in Multi-attribute Decision Models. *Journal of Multi-Criteria Decision Analysis* 11, pp. 291-303, Wiley.

Roy, B. and Vanderpooten, D. (1996). The European School of MCDA: Emergence, Basic Features, and Current Works. *Journal of Multi-Criteria Decision Analysis* 5, pp. 22-37, Wiley.

Saaty, T. L. (1977). Scenarios and priorities in transport planning: Application to the Sudan. *Transport Research* 11, pp. 343-350, Pergamon Press.

Saaty, T. L. (2001). *Decision Making for Leaders: The Analytical Hierarchy Process for Decisions in a Complex World*. The Analytical Hierarchy Process Series, Vol. 2, Pittsburgh, United States of America.

Salling, K. B. (2008). Assessment of Transport Projects – Risk Analysis and Decision Support. Ph.D. thesis, Department of Transport, Technical University of Denmark.

Stewart, T. J. (1992). A Critical Survey on the Status of Multiple Criteria Decision Making Theory and Practice. *Omega* 20, pp. 569-586, Pergamon Press.

Stewart, T. J. (1993). Use of Piecewise Linear Value Functions in Interactive Multicriteria Decision Support: A Monte Carlo Study. *Management Science* 39, pp. 1369-1381, Informs.

Stewart, T. J. (1996). Robustness of Additive Value Function Methods in MCDM. *Journal of Multi-Criteria Decision Analysis* 5, pp. 301-309, Wiley.

Vaidya, O. S. and Kumar, S. (2006). Analytic Hierarchy Process: An Overview of Applications. *European Journal of Operational Research* 169, pp. 1-29, Elsevier.

van den Honert, R.C. and Lootsma, F.A. (2000). Assessing the quality of negotiated proposals using the REMBRANDT system. *European Journal of Operational Research* 120, pp. 162-173, Elsevier.

von Winterfeldt, D. and Edwards, W. (1986). *Decision Analysis and Behavioural Research*. Cambridge University Press, New York, United States of America.

von Winterfeldt, D. and Edwards, W. (2007). Defining a Decision Analytic Structure. In: Edwards, W., Miles, R. F. and von Winterfeldt, D. (Eds.). *Advances in Decision Analysis – From Foundations to Applications* (pp. 81-104). Cambridge University Press, New York, United States of America.

Watson, S. R. and Buede, D. M. (1987). *Decision Synthesis: The Principles and Practise of Decision Analysis.* Cambridge University Press New York, United States of America.

World Bank (2010). Cost-Benefit Analysis in World Bank Projects. Report from the Independent Evaluation Group (IEG), Washington, D.C, United States of America, ISBN 978-0-8213-8578-4.

Paper 1

Composite decision support by combining cost-benefit and multi-criteria decision analysis

Michael Bruhn Barfod
Department of Transport, Technical University of Denmark

Kim Bang Salling
Department of Transport, Technical University of Denmark

Steen Leleur
Department of Transport, Technical University of Denmark

Paper published in *Decision Support Systems* 51, Issue 1, pp. 167-175 (Elsevier), 2011 – ISI-indexed

Keywords: Decision support systems, decision analysis, cost-benefit analysis, multi-criteria decision analysis

Abstract

This paper concerns composite decision support based on combining cost-benefit analysis (CBA) with multi-criteria decision analysis (MCDA) for the assessment of economic as well as strategic impacts within transport projects. Specifically a composite model for assessment (COSIMA) is presented as a decision support system (DSS). This COSIMA DSS ensures that the assessment is conducted in a systematic, transparent and explicit way. The modelling principles presented are illuminated with a case study concerning a complex decision problem. The outcome demonstrates the approach as a valuable DSS, and it is concluded that appraisals of large transport projects can be effectively supported using a combination of CBA and MCDA. Finally, perspectives of the future modelling work are given.

1. Introduction

Decision support systems (DSS) are widely applied to assist decision-makers with the difficult task of identifying the best solution to a given problem. This paper concerns composite decision support based on combining cost-benefit analysis (CBA) with multi-criteria decision analysis (MCDA) for the assessment of economic as well as strategic impacts within transport projects. Specifically a composite model for assessment (COSIMA) is presented as a decision support system (DSS). This COSIMA DSS ensures that the assessment is conducted in a systematic, transparent and explicit way. The approach is presented through a case study concerning alternatives for the construction of a new fixed link between the city of Frederikssund and the peninsula of Hornsherred on the north-eastern part of Zealand, Denmark (Barfod, 2006).

The most common methodology applied so far to the evaluation of transport systems has been conventional CBA (Janic, 2003), which supported by traffic- and impact model calculations provides the decision-makers with a monetary assessment of the project's feasibility. A socio-economic analysis is in this respect a further development of the traditional CBA capturing the economic value of social benefits by translating social objectives into financial measures of benefits (Wright et al., 2009). Internationally seen there has been a growing awareness over the recent years that besides the social costs and benefits associated with transport other impacts that are more difficult to monetise should also have influence on the decision making process. This is in many developed countries realised in the transport planning, which takes into account a wide range of impacts of also a strategic character (van Exel et al., 2002). Accordingly, appraisal methodologies are undergoing substantial changes in order to deal with the developments (Vickerman, 2000) that are varying from country to country and leading to different approaches (Banister and Berechman, 2000). It is, however, commonly agreed that the final decision making concerning transport infrastructure projects in many cases will depend on other aspects besides the monetary ones assessed in a socio-economic analysis. Nevertheless, an assessment framework such as the Danish one (DMT, 2003) does not provide any specific guidelines on how to include the strategic impacts; it merely suggests describing the impacts verbally and keeping them in mind during the decision process.

A coherent, well-structured, flexible, straight forward evaluation method, taking into account all the requirements of a transport infrastructure project is for this reason required. An appropriate ex-ante evaluation method for such projects can be based on MCDA (Tsamboulas, 2007; Vreeker et al., 2002), which in most cases can be combined with a CBA (Leleur, 2000). Scanning the literature (Goodwin and Wright, 2009; Belton and Stewart, 2002; Keeney and Raiffa, 1993; von Winterfeldt and Edwards, 1986) it is found that the use of MCDA in the decision process usually provides some or all of the following features:

1. Improvement of the satisfaction with the decision process
2. Improvement of the quality of the decision itself
3. Increased productivity of the decision-makers

MCDA can in this respect be seen as a tool for appraisal of different alternatives, when several points of view and priorities are taken into account to produce a common output. Hence, it is very useful during the formulation of a DSS designed to deal with complex issues. The literature on Decision Support Systems is extensive, providing a sound basis for the methodologies employed and the mathematics involved. Moreover, there are numerous systems covering several disciplines, policy contexts and users' needs for specific application environments (Salling et al., 2007; Tsamboulas and Mikroudis, 2006; Janic, 2003). The use of DSS for solving MCDA problems has among others been treated by Chen et al. (2008) and Larichev et al. (2002), where it is shown that a DSS can effectively support a decision making process making use of appropriate MCDA methodologies.

Earlier research on composite DSS within transport planning has mainly concentrated on incorporating the CBA in the MCDA. Here the European Commission's fourth framework project EUNET (2001), which has developed a methodology dealing with the combination of CBA and MCDA, can be mentioned. The EUNET framework applies scores to the investment criterion, e.g. the benefit cost rates (B/C-rate), thus, it treats the rates as any other criterion in the MCDA. Exactly which criteria to include in the framework is a matter of judgment depending, among other factors, on the reliability of the data and the preferences stated by the decision-makers and/or stakeholders in the decision process. Another similar inclusive approach is proposed in Sayers et al. (2003) for transport project appraisal in the UK. Different methodological frameworks are used varying from country to country; however, it is roughly possible to divide them into two main categories: CBA-based and MCDA-based frameworks. Among the CBA-based frameworks the Danish and German can be mentioned, while e.g. French and Dutch frameworks are based on the use of MCDA (for further information about European frameworks see e.g. Banister and Berechman (2000) and EUNET (2001)). Reviews of transport appraisal methodologies and their premises and results can be found in for example Hayashi and Morisugi (2000) and Mackie and Preston (1998) listing also various sources of error and bias in them.

The COSIMA DSS presented and discussed in this paper provides a theoretical and practical methodology for adding non-monetary MCDA-criteria to the monetary CBA-impacts. Unlike previous attempts this DSS is based on the argument that the MCDA-criteria can be added to the CBA-impacts – if value functions can be computed for the MCDA-criteria using a weighting procedure describing the importance of each criterion. Hence, the COSIMA approach is based on the theoretical valid and widely used methodology of additive value functions (see e.g. Keeney and Raiffa (1993) or von Winterfeldt and Edwards (1986)). The value function scores are in the presented COSIMA set-up derived via direct rating using pair wise comparisons (Belton and Stewart, 2002) and the criteria weights are derived by applying rankings (Roberts and Goodwin, 2002) in this respect drawing on the Analytic Hierarchy Process (AHP) and the Simple Multi-Attribute Rating Technique Exploiting Ranks (SMARTER) respectively. The total benefits from the CBA are used to determine shadow prices for the MCDA-criteria; this is based on a trade-off between the CBA and MCDA, which creates a total rate of return (TRR). The last part requires converting strategic non-quantitative criteria into monetary values (Tsamboulas and Mikroudis (2000), which will be discussed further in this paper.

Additive value functions and the AHP are well established and widely applied methods for MCDA (Tsamboulas, 2007; Belton and Stewart, 2002; Saaty, 2001; Hwang and Yoon, 1995). Consequently, they seem appropriate to use in the proposed DSS. Moreover, in terms of transparency, the additive model appears most favourable, since it is considered to be able to cope with almost any problem (Jiménez et al., 2003; Belton and Stewart, 2002). It is, however, commonly known that the main drawback of these methods is the assignment of criteria weights, since individuals are determining these weights. On the other hand, the performance (scores) of alternatives for each criterion is determined objectively, even if artificial scales are used for non-quantifiable criteria. However, the AHP contributes to overcoming this disadvantage by deriving weights in a quasi-independent manner, using pair wise comparisons that make it difficult to promote open biases towards specific criteria. Thus, AHP is a common method used for prioritisation when having a wide variety of choices. More specifically, with regard to the application of the DSS for the case study, the group that assigned the weights was composed of key stakeholders involved in the project.

With reference to the previous work on composite DSS conducted by other researchers this paper deals with three main research questions: Can comprehensive appraisals taking into account both monetary impacts and non-monetary criteria of a decision problem be operationalised to a DSS that can inform the users in terms of both interaction and interpretation of the results? Can a valid guidance be formulated for adding the non-monetary criteria of the MCDA to the monetary impacts of the CBA in the DSS? And finally, can a set of appropriate guidelines for inclusion of non-monetary criteria in transport planning be set out?

The paper is organised as follows. After this introduction the principles for the modelling approach of the COSIMA DSS is presented. Following, the case study regarding the appraisal of the Danish Hornsherred case is presented and the COSIMA DSS is applied in terms of a comprehensive assessment by incorporating respectively a CBA and a MCDA leading to a composite result. Finally, a conclusion is drawn and perspectives for the future modelling work are given.

2. Modelling approach

As mentioned in the introduction the appraisal and planning of transport infrastructure projects should be based on all relevant impacts, which are depending on the type and size of the project viewed upon. Some of these impacts can be assessed monetarily and are thereby possible to include in a socio-economic CBA in a decision model as e.g. the CBA-DK model (Salling, 2008). However, no common guidelines exist for the assessment of impacts such as urban development, landscape, etc. that hold a potential for improving actual decision support from the assessment if they are treated properly. These non-monetary or strategic impacts should instead be assessed using a MCDA and are hence denominated MCDA-criteria. The idea behind composite modelling assessment is to extend the CBA into a more comprehensive type of appraisal as often demanded by decision-makers by including these 'missing' decision criteria of relevance for the actual assessment task.

2.1 Principles for the COSIMA DSS

The COSIMA DSS consists of a CBA-part and a MCDA-part and the result of the COSIMA analysis is expressed as a total value (*TV*) based on both parts. This model set-up emphasizes that the MCDA-part should be truly additive to the CBA-part. For this reason a project alternative, A_k, is better represented towards the decision making by the $TV(A_k)$ rather than by the present value from the CBA (the sum of all benefits and disbenefits) – here denominated $CBA(A_k)$. The principle in COSIMA can be expressed by (1), where $MCDA(A_k)$ represents a term which is added to $CBA(A_k)$ (Salling et al., 2007):

$$TV(A_k) = CBA(A_k) + MCDA(A_k) \tag{1}$$

The assessment principles in the MCDA-part are based on decision-maker involvement. This is not the case in the conventional CBA. This circumstance justifies the MCDA denomination as the part is based on subjective assessments. It can be noted on the basis of (1) that a situation, where $CBA(A_k)$ is equal to or smaller than the investment costs (C_k), is non-profitable seen from a socio-economic point of view (i.e. $CBA(A_k) \leq C_k$). However, the investment can still be justified by the wider COSIMA analysis if the total value of A_k is larger than the investment costs ($TV(A_k) > C_k$). This can also be expressed via the total rate of return (*TRR*) calculated as the total value, $TV(A_k)$, divided by the investment cost, C_k, which gives $TRR(A_k) > 1$, indicating a total rate with regard to the attractiveness of alternative k. This leads to (2) below:

$$TRR(A_k) = \frac{TV(A_k)}{C_k} = \frac{1}{C_k} \cdot \left(\sum_{i=1}^{I} V_i(X_{ik}) + \alpha \cdot \left[\sum_{j=1}^{J} w_j \cdot VF_j(Y_{jk}) \right] \right) \tag{2}$$

where

$$\sum_{j=1}^{J} w_j = 1 \text{ and } 0 < w_j < 1$$

- A_k is alternative k
- C_k are the total costs or expenses of alternative k
- X_{ik} is the quantity of the CBA impact i for alternative k
- $V_i(X_{ik})$ is the value in monetary units for the CBA impact i for alternative k
- α is an indicator that expresses the balance between the CBA and MCDA parts in the model
- w_j is a weight which reflects the importance of MCDA criterion j
- Y_{jk} is a parameter value for MCDA criterion j for alternative k
- $VF_j(Y_{jk})$ is a value-function score for MCDA criterion j for alternative k

The general COSIMA approach is presented by (1) and (2). (2) can be specified into a CBA if sufficient knowledge about the criteria to be assessed in the MCDA part is available (e.g. a unit price for at least one criterion and true importance weights for all criteria). This would for example be the case if a conventional CBA is carried out and afterwards supplemented with some extra criteria with satisfactory unit prices specified fully by impact models. However, this will most often not be possible as the impacts handled in the MCDA part in general are based on non-empirical knowledge and often cannot be determined by impact models or even assigned with a unit price. Hence, the purpose of COSIMA is to handle such a situation in a comprehensive and transparent way. This ensured through the determination of appropriate values for α and w_j for the J MCDA-criteria and appropriate value-function scores $VF_j(Y_{jk})$. The latter supplements of (2) the determination of $V_i(X_{ik})$ which can be derived in accordance with a socio-economic manual relevant for the actual assessment case (DMT, 2003).

2.2 Calibration of the COSIMA DSS

Regarding the α-indicator, that expresses the balance between the CBA and MCDA parts in the model set-up, it should be noted that the CBA calculation remains unchanged in COSIMA, but that different α-values will change the MCDA's influence on the TRR. In practice it has been found convenient to express α based on a MCDA%, which reflects the relative weight of the MCDA-part compared to the CBA-part. The value of $\alpha = \alpha(\text{MCDA\%})$ is then set by determining MCDA% $= 100 \cdot \sum_j B_j / [\sum_i B_i + \sum_j B_j]$, where $B_i = \sum_{\kappa \in K}(b_{ik})$ and $B_j = \sum_{\kappa \in K}(b_{jk})$ represent the value elements for the individual CBA-impact i and MCDA-criterion j summed over the κ alternatives (the alternatives A_κ chosen for calibration of the model). Thus $\sum_i B_i$ and $\sum_j B_j$ are summations of the I CBA-impacts and the J MCDA-criteria, and B_i and B_j the results of the b_{ik} and b_{jk} summations of the alternatives, where some if not all are selected for the model calibration (Leleur, 2008).

The calculations in the COSIMA DSS use a parameter for the calibration named UP_j which functions as a shadow price per index value for each of the J MCDA-criteria in order to produce the b_{jk} values. These benefit values obtained are determined by $b_{jk} = VF_j(Y_{jk}) \cdot UP_j$ where the shadow price, UP_j, is a function of the α-indicator (MCDA%), the criteria weights (w_j), the sum of benefits from the CBA for the alternatives used for calibration $(\sum_i \sum_{\kappa \epsilon K}(b_{i\kappa}))$ and the sum of VF-scores for the alternatives used for calibration $(\sum_\kappa VF_j(Y_{jk}))$, see (1) and (2). In the procedure $\alpha($MCDA%$)$ and w_j determine a fraction of $\sum_i \sum_{\kappa \epsilon K}(b_{i\kappa})$ that by unit scaling leads to the J unit prices that are used for calculating the $TRR(A_k)$. It should be noted that TRR-values are also calculated for alternatives not used in the calibration set. Changes in the set of alternatives A_κ behind the calibration will influence the UP_j values and thereby the TRR-values. This pool dependence is of great interest for the decision analyst who is formulating the model set-up. The alternatives should in this respect be scrutinised so that the calibration pool only consist of alternatives that are possible solutions (Ibid).

Finally, it is important to note that the TRR-value due to the theoretical differences between CBA and MCDA has no economic argument like e.g. the B/C rate, see Figure 1. However, the TRR makes it possible to incorporate and hereby retain the information from e.g. the B/C rate in its original form. Additionally, the TRR delivers information concerning the expression of the alternatives performance in relation to the MCDA-criteria – all included in one single rate. Hence, the COSIMA approach provides the decision-makers with a composite measure of attractiveness.

Figure 1 depicts an overview of the methodological approach resulting in the TRR incorporating the CBA (described in section 4), the MCDA applying the AHP and SMARTER techniques (described in section 5) and the summation in COSIMA using the MCDA % (described in section 6).

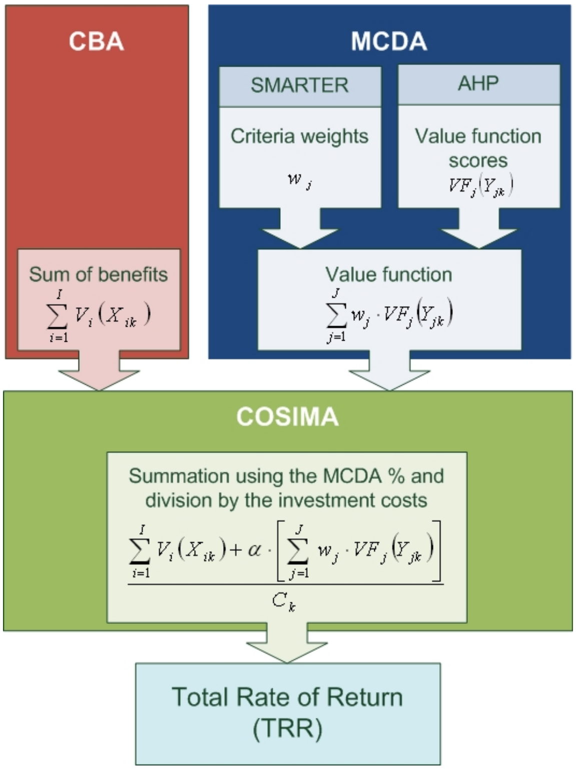


Figure 1. Overview of the steps in the COSIMA DSS

3. Case presentation

The case study considered concerns the city of Frederikssund which is situated at Roskilde Fjord in the northern part of Zealand, Denmark. The fjord is separating a peninsula, Hornsherred, from Frederikssund and the rest of the Greater Copenhagen area. Currently, the only connection across the fjord is a bridge featuring only one lane in each direction. This is creating a huge bottleneck for the traffic in the area around the city of Frederikssund. Moreover, the Danish government has current plans for the construction of a new motorway between Copenhagen and Frederikssund; this will only lead to a further increase of the traffic problems in the area. Several preliminary studies with the purpose of finding a solution to the traffic problems have been conducted by the municipality of Frederikssund in cooperation with the Danish Ministry of Transport (Barfod, 2006). According to Barfod (2006) only four alternative solutions seem realistic for relieving the current traffic situation in Frederikssund:

1. An upgrade of the existing road through Frederikssund and the construction of a new bridge parallel to the old bridge
2. A new high-level bridge and a new by-pass road south of Frederikssund
3. A short tunnel with embankments and a new by-pass road south of Frederikssund
4. A long tunnel without embankments and a new by-pass road south of Frederikssund

According to the characteristics of the above mentioned alternatives different impacts will be derived from each alternative implying different investment costs and layouts. The primary stakeholders behind the project, the Region and the municipalities in the area, have formulated the goal and objective of the project as follows (Barfod, 2006):

> *Improve the traffic flow across the fjord … the solution should take great considerations as concerns the environment in the form of traffic derived consequences (e.g. noise and air pollution) and furthermore consequences derived from the construction itself (e.g. nature and landscape).*

This statement calls for a broader type of appraisal than a conventional CBA, which as mentioned in Denmark only includes the first type of consequences, however, supplemented with a verbal description of the last type. For this reason the comprehensive approach of the COSIMA DSS embracing both monetary and more strategic consequences is applied to the case study in order to produce informative decision support to the decision-makers. In order to make the appraisal of the alternatives as comprehensive as possible, representatives for key players involved in the decision process are gathered to systematically discuss and analyse the issues at a so-called decision conference as described by Phillips (1984; 2006). The objective of such a decision conference is to constructively deal with the conflicting issues at hand so that a common understanding of the issue can be achieved (Mustajoki et al., 2007). The COSIMA DSS is consequently used to model the viewpoints of the participants and to evaluate the alternatives in an auditable manner.

4. Cost-benefit analysis

The first step in the COSIMA analysis is to conduct a socio-economic CBA (to derive: $V_i(X_{ik})$ from (2)). This CBA is carried out in a model named the CBA-DK model (Salling, 2008) in accordance with the Danish Manual for Socio-Economic Appraisal (DMT, 2003). Thus, the CBA includes an assessment of the principal items: Construction and maintenance costs, travel time savings and other user benefits, accident savings and other external effects, taxation, scrap value, and finally tax distortion. A traffic- and impact model is set up in order to calculate the impacts derived from each project alternative and the construction and maintenance costs are estimated (Barfod, 2006). All impacts are then entered into the CBA-DK model, where forecasting is added according to assumptions about the future development in traffic. The various economic parameters necessary for conducting the CBA are set in accordance with Danish standards (DMT, 2003) and will not be treated further in this paper. Finally, the feasibility of the alternatives is described by three different investment criteria in the model (see Figure 2 and Table 1): The benefit cost rate (B/C-rate), the internal rate of return (IRR) and the net present value (NPV). For elaborating description of the investment criteria see e.g. Leleur (2000). The results for all the four alternatives are shown in Table 1.

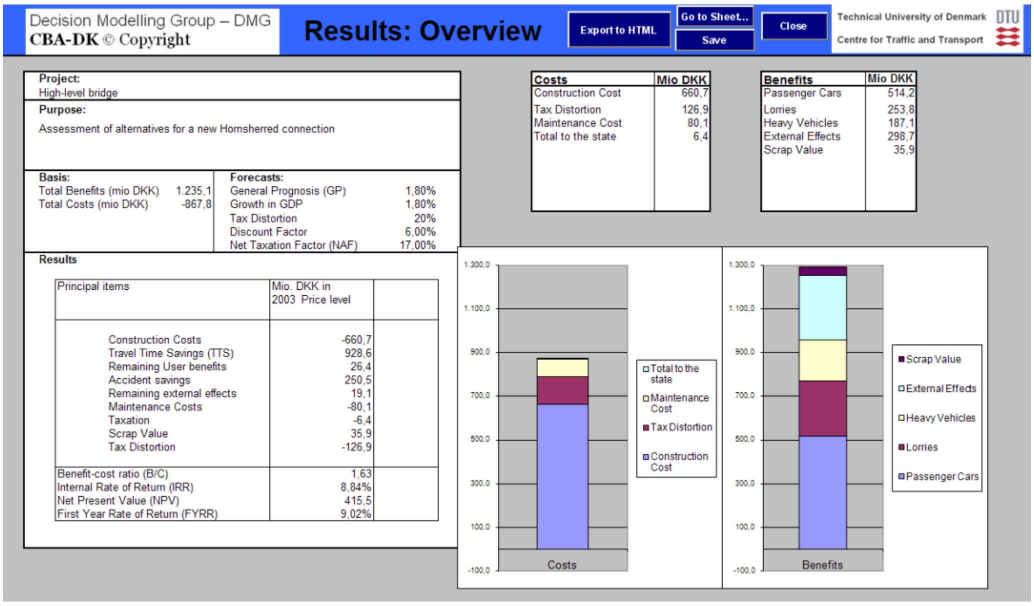


Figure 2. Results of the cost-benefit analysis for the High-level bridge alternative presented in the CBA-DK model

Figure 2 depicts how the results are presented in the CBA-DK model. The principal items of a fixed CBA-DK model run are listed on the left hand side, and the two columns on the right hand side show the size of the costs and the benefits in the same absolute scale. The results for all the four alternatives are shown in table 1.

Table 1. Investment criteria and the sum of benefits ($V_i(X_{ik})$) for the four alternatives concerning the case

	High-level bridge	Short tunnel	Long tunnel	Upgrade
B/C-rate	**1.63**	0.99	0.25	0.36
IRR [%]	**8.84**	5.95	2.06	2.76
NPV [m DKK]	**415.5**	-10.2	-1,869.3	-235.6
$V_i(X_{ik})$ [m DKK]	1076	965	607	133

The CBA results clearly show that a high-level bridge is the only feasible alternative if the decision is to be based solely on monetary impacts. The three other alternatives are not feasible according to the investment criteria; the short and the long tunnels because of their high construction costs and the upgrade because of its less significant user benefits.

5. Multi-criteria decision analysis (MCDA)

The use of MCDA is aimed at supporting decision-makers who are faced with numerous and conflicting choices (Lootsma, 1999). Unlike methods like CBA that assume the availability of empirical data, data in MCDA are derived or interpreted subjectively as indicators of the strength of the decision-makers preferences. These preferences differ from decision-maker to decision-maker; hence, the outcome of the analysis depends on who are making the assessments and what their goals and preferences are. Since MCDA involves a certain element of subjectiveness, the morals and ethics of the decision-makers implementing MCDA play a significant part in the accuracy and fairness of MCDA's conclusions. The transparency of the assessment is in this respect very important when one is making a decision that seriously impacts on other people. Generally, the different methods that exist for conducting MCDA have been designed in order to designate a preferred alternative, to classify the alternatives in a small number of categories, and/or to rank the alternatives in a subjective order of preference. The second step of the COSIMA analysis, thus, comprises a MCDA in order to determine scores for the alternatives and weights for the criteria, i.e. $VF_j(Y_{jk})$ and w_j from (2).

Applying creative techniques such as brainstorm at the decision conference as described by Jeppesen (2010) the respondent group decided to include four different MCDA-criteria for the case study, covering the ‘missing’ effects of the CBA. Special attention was made in the definition of the criteria in order to avoid double counting. The criteria definitions are depicted in Table 2.

Table 2. The criteria to be assessed by the MCDA

	Definition
Accessibility	The criterion ranges from local accessibility through regional accessibility to public accessibility and favours alternatives that contribute to improve the overall accessibility in the transport network.
Urban development	The criterion favours alternatives that contribute to develop the existing parts of the city considered plus the opportunity to expand the city and develop new parts.
Landscape	The criterion covers the alternatives visual impact on the landscape and favours those alternatives, which have the least negative impact on this plus on recreational areas and areas worthy of preservation.
Environment	The criterion covers the environmental issues that are not treated in the CBA, i.e. maritime conditions in the fjord plus plant and animal life in and around the fjord.

5.1 Scoring of alternatives

In order of determining the impact of the alternatives within the MCDA-criteria (the value function scores $VF_j(Y_{jk})$ from (2)) an appropriate assessment technique (MCDA method) has to be chosen. Several assessment techniques are available for the purpose of determining the value function scores ($VF_j(Y_{jk})$). However, as one of the purposes set out for the COSIMA DSS is to assure to be transparent and easily understandable for decision-makers the well established and widely applied pair wise comparison technique AHP (Analytic Hierarchy Process) by Saaty (1977) is used. It should be noted, that even though the AHP technique is considered to be transparent and appropriate for the current case study, other more or less complicated case studies may call for other techniques (see e.g. Goodwin and Wright (2009), Belton and Stewart (2002), and Olson et al. (1995) for other MCDA methods).

Using the AHP technique, the decision-makers are involved in direct ratings via pair wise comparisons (Belton and Stewart, 2002; Saaty, 2001) of the alternatives within each of the criteria. For each comparison the decision-makers have to state the strength of their preference for one alternative over another according to the semantic scale that spans from equal preference to absolute preference (1 to 9 on the numerical scale) (Saaty, 1977). Next, the scale values obtained by the pair wise comparisons are for each criterion implemented in a comparison matrix, and normalised scores (AHP scores) for the alternatives are derived using the geometric mean method (Hwang and Yoon, 1995; Barzilai et al., 1987). These normalised scores are computed into value-function (VF) scores utilizing a local scale from 0-100, where the score 0 is describing the worst performing alternative and the score 100 is describing the best performing alternative (Belton and Stewart, 2002). All other alternatives will receive linear intermediate scores relative to the two end points. It is noted that the use of a local scale limits the appraisal to concern only the relationship between the already identified alternatives; no absolute measure of their performance is obtained. A local scale is considered to be a useful solution when only dealing with alternatives for one project, i.e. dealing with a closed system. The local scale would, however, not be adequate if it should be possible to include more alternatives at a later stage in the appraisal. In such a case it would be necessary to revise the scale or use a global scale taking the extreme endpoints into account (for more information about local versus global scales see e.g. Belton and Stewart (2002)). The COSIMA DSS is customised to the specific problem at hand and thereby assumes that no other alternatives than those identified at the preliminary stage, will be relevant for the appraisal, hence the local scale is appropriate to use.

As a result of the choice of the AHP technique the participants at the decision conference were faced with full pair wise comparisons of the four alternatives within the four criteria comprising a total of 24 comparisons. In order to assure the validity and reliability of the comparisons, time was allocated for a thorough discussion of each comparison and the rationale were recorded in an assessment protocol. Efforts were in this respect made for the participants to reach consensus on each of the comparisons before moving on to the next. In the cases where it was not possible to agree upon the comparisons the different viewpoints were noted with a view to a later sensitivity analysis if felt needed by the participants. In addition to these efforts a consistency check of the comparisons were made and the participants were notified and asked to revise one or more comparisons if the inconsistency index exceeded 0.1 (Goodwin and Wright, 2009).

Next, the AHP-scores derived from the input of the participants are computed into VF-scores. The VF-scores are for each of the four project alternatives shown in Table 3. Considering these scores special attention should be paid if nearly identical AHP-scores are derived and succeedingly transformed into very varying VF-scores. If this is the case, the criterion from the sample should be assigned with a low weight or maybe even omitted as it does not contribute to the segregation between alternatives and consequently it will not be relevant to include in the appraisal. The decision analyst should after the completion of the pair wise comparisons study the VF-scores, perform a check-up with regard to the above and notify the participants if the issue above is relevant.

Table 3. Value function scores describing the alternatives performance within the criteria

	High-level bridge	Short tunnel	Long tunnel	Upgrade
Accessibility	100	100	100	0
Urban development	100	100	100	0
Landscape	0	7	34	100
Environment	8	0	46	100

Table 3 depicts the VF-scores for the four alternatives within the four types of MCDA-criteria. It is noted that three of the alternatives, i.e. the High-level bridge, the Short tunnel and the Long tunnel, have identical VF-scores for “accessibility” and “urban development” namely 100 while the Upgrade’s VF-scores are 0 for both criteria. This is due to the former’s identical alignments, which have much larger impact within the two criteria than the latter. Hence, the large span between 0 and 100 seems reasonable within the criteria. Within the “landscape” and “environment” criteria the alternatives differentiate more between each other and further investigation is not needed. Clearly, the Long tunnel alternative has the overall best performance within the four criteria. It is for this reason the most attractive alternative if the importance of each of the four criteria is weighted equally. However, the decision-makers would most often feel that some criteria are more important than other. Thus, a weighting procedure describing the importance of each criterion is assigned.

5.2 Weighting of criteria

For the determination of the criteria weights (w_j) the SMARTER (Simple Multi-Attribute Rating Technique Exploiting Ranks) technique (Goodwin and Wright, 2009) using ROD (Rank Order Distribution) weights (Roberts and Goodwin, 2002) is applied to the COSIMA DSS. The ROD-weights are surrogate weights which provide an approximation to unrestricted original weights. Surrogate weights based on rankings have been proposed as a method for avoiding the difficulties associated with the elicitation of weights in MCDA (Belton and Stewart, 2002). The decision making process is thereby simplified as the technique only requires an importance ranking of the four MCDA-criteria. Thus, no specification of the weights is needed from the decision-makers as these are determined by the ROD technique and assigned to the criteria according to the ranking. Hence, the participants at the decision conference were faced with the task of ranking the criteria in order of importance. Different viewpoints were expressed by the participants; however, it was possible through discussion to reach consensus in the group. The ranking agreed upon is in correspondence with the SMARTER technique assigned with the ROD-weights as shown in Table 4.

Table 4. ROD weights assigned to the criteria according to the level of importance

Rank	Criteria	ROD-weight
1	Accessibility	0.42
2	Environment	0.30
3	Landscape	0.19
4	Urban development	0.09

Determining the weights requires much responsibility and expertise from the decision-makers as the weights have considerable influence on the results of the assessment. Using the SMARTER technique applying ROD weights instead of using pair wise comparisons from AHP to determine the weights has for this purpose been chosen in order to make the interaction with the decision-makers around weights simpler. It is assumed that the setting of weights for criteria is more subjective than the scoring of the project attributes based on pair wise comparisons. For the latter the consistency AHP check has the purpose of securing certain validity for the scores.

6. Combining CBA and MCDA

The last parameter in (2) – the α indicator – is determined as a balance (weight) between the CBA and the MCDA and expressed by the MCDA% as described in section 2.2. The calibration of the model was made using all four alternatives as all were regarded to be serious contenders for the final choice; the High-level bridge due to its performance within both the CBA and MCDA, and the three other alternatives due to their performance within the MCDA. As noted, the CBA calculation result remains unchanged at all stages in the composite appraisal, but different values of α (the MCDA%) will change the weight of the MCDA on the TRR. As depicted in (2) the MCDA% will always be less than 100 as the CBA outcome always will have influence on the TRR i.e. the CBA result cannot be omitted from the appraisal.

Figure 3 discloses the result covering the agreed ranking of the MCDA-criteria. Note that MCDA% values larger than 80 are not shown on the figure as no changes in the ranking between the alternatives take place based on these values.

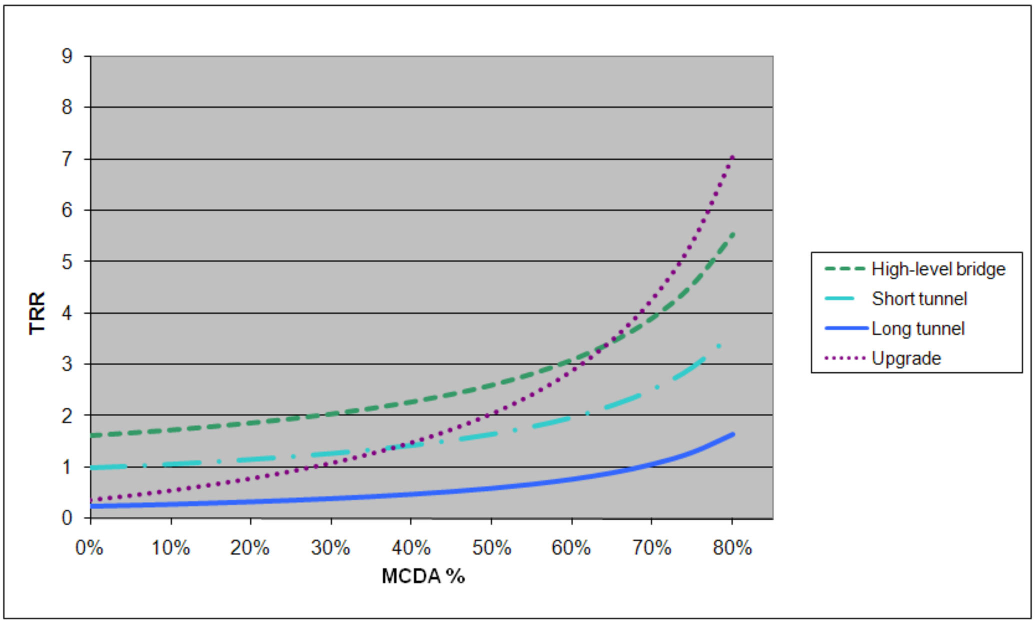


Figure 3. TRR values for the four alternatives as a function of the MCDA%

Figure 3 depicts that the TRR-values increase for all four alternatives as the MCDA adds a higher importance. However, it can be seen that the TRR for the upgrade of the existing connection is increasing more rapidly than the other alternatives which is due to lower construction costs for this alternative, compared with the three others. The increase, however, is also very dependent on which MCDA criteria that are considered the most important ones. The ranking of criteria in Table 4 shows accessibility and environment as the most important criteria. As a result of this the alternatives with high scores on these criteria, will also have the largest increase in their TRR-value and vice versa. The TRR-values point out different alternatives as the most attractive depending on which MCDA% that is considered. However, it is not always beneficial to let the decision-makers base their decision on an interval result, e.g. from 0 to 80 %. In most cases the decision-makers should agree upon a specific MCDA% or a short interval (e.g. 30 – 50 % MCDA) before deciding about the project. Practical experience so far points to MCDA% values in the range between 20 to 30 % (Barfod, 2006). Furthermore, it seems that high MCDA% values are most likely to be adopted when appraising larger and more complex strategic infrastructure projects e.g. the Fehmarn Belt fixed link between Germany and Denmark. The MCDA% to base the decision on will also vary depending on society's economy, current political tendencies and the type of project being appraised.

The participants at the decision conference were asked to express their preferences with regard to the CBA/MCDA weighting and there was agreement that CBA was the most important part of the appraisal as the project is not regarded to be one of the large strategic infrastructure projects mentioned before. After a discussion a MCDA% set to be 30 was chosen, and with this CBA/MCDA balance applied the high-level bridge is still clearly the most attractive alternative. Thus, this alternative appears as the most robust choice based on the comprehensive appraisal.

7. Results and discussion

Summarising the calculations and the process concerning the case it is noted that the analysis is based on the use of CBA and MCDA. The CBA produces results that can be measured in a monetary unit – here million Danish Kroner (m DKK). The MCDA on the other hand produces results that by comparison with the CBA results can be calibrated to 'assessment m DKK'. In order to obtain a total value for an examined alternative m DKK and 'assessment m DKK' is added. This mix between m DKK and the fictitious 'assessment m DKK' is expressed by the unit 'attractiveness m DKK'. The result can also be presented as a total rate of return (TRR), where the result in 'attractiveness m DKK' is divided by the investment costs, see Table 5. In this context it should be noted that the process based on input (scoring of alternatives, determination of criteria-weights and balancing the CBA and MCDA) makes it possible to provide a transparent evaluation process involving the decision-makers.

Thus, the result of the COSIMA analysis is that using decision-maker involvement it is possible to apply values to the MCDA-criteria which are comparable to the monetary values from the CBA. The results depicted in Table 5 indicate the 'gain' by choosing an alternative which performs well within the MCDA instead of the alternative which performs the worst. In strategic terms the decision-makers in the case study would achieve most from the investment by choosing the Long tunnel alternative (MCDA alone). However, overall (CBA + MCDA) the High-level bridge alternative will continue to be the most attractive.

Table 5. Results of the composite COSIMA analysis using a MCDA% set to be 30

	High-level bridge	**Short tunnel**	**Long tunnel**	**Upgrade**	**Method**	**Unit**
Investment costs	661	975	2477	369	CBA	m DKK
Total benefits	1076	965	607	133	CBA	m DKK
B/C rate	**1.63**	**0.99**	**0.25**	**0.36**	CBA	
Accessibility	*165*	*165*	*165*	*0*	MCDA	Assessment m DKK
Urban development	*101*	*101*	*101*	*0*	MCDA	Assessment m DKK
Landscape	*0*	*8*	*42*	*122*	MCDA	Assessment m DKK
Environment	*11*	*0*	*66*	*144*	MCDA	Assessment m DKK
Total MCDA	*277*	*274*	*374*	*266*	MCDA	Assessment m DKK
Total value	***1353***	***1239***	***981***	***399***	CBA+MCDA	m DKK + ’Assessment m DKK’ = 'Attractiveness m DKK'
Total rate	***2.05***	***1.27***	***0.40***	***1.08***		

As mentioned in section 1 the existing assessment framework in Denmark does not attempt to incorporate the strategic issues (the MCDA-criteria) of a decision problem into appraisals of transport infrastructure projects. Other frameworks, such as the EUNET framework (EUNET, 2001), incorporate the CBA results as a criterion in the MCDA and the result is expressed in form of a relative rate. Using the COSIMA DSS the decision-makers are provided with a result that contains a level of information which comprises both the CBA and MCDA expressed in a more easy accessible way. Generally, decision-makers are used to make decisions on the basis of a B/C rate and are hence comfortable with this type of expression. The new feature in the COSIMA DSS is that the MCDA part is converted to the same scale as the CBA part providing the decision-makers with an indication of the value of the strategic issues based on their own preferences expressed as a total rate of return. The TRR result will most likely vary based on who is stating the preferences, however, by assuring diversity in the assessment group the result becomes valid to a wide audience.

A downside of using a pair wise comparison technique such as AHP is the number of comparisons that the respondent group has to make. In the case study addressed in this paper only four alternatives are to be assessed within four criteria leading to 24 comparisons, but if just one extra alternative is added to the appraisal the number of comparisons will be 40, if another is added the number is 60 etc. Hence, the more alternatives in an assessment the more inappropriate the pair wise comparison technique becomes. If there are too many comparisons to be made the respondent group tends to get tired and make comparisons of a lower quality as their will to discuss fades; the comparisons can then tend to be taken as more or less an average of the groups’ viewpoints. This will influence the rest of the appraisal and generate poor results. In order for this not to be an issue another assessment technique should be chosen if the number of comparisons exceeds what seems reasonable to manage within the given time frame of the assessment task. From testing the methodology at a number of occasions it has been found that the maximum number of comparisons demanded from the respondent group should be less than 50 if the time frame available is only one day or less. In one test case this limit was reached as the respondent group had to assess four alternatives within 8 criteria leading to 48 comparisons in a half day meeting.

An important issue to address when making the final conversion of the MCDA part to the CBA part is the importance of the criteria weights. The weights are directly linked to the shadow prices assigned to the criteria in the COSIMA-DSS and changes in the weights will thus have a large influence on the final outcome of the analysis. The criteria weights can be seen as the most subjective part of the MCDA assessment and will differ dependent on who is setting them. For this reason sensitivity analysis should be conducted testing different weight sets in order to see how changes will affect the investment decision to be taken.

8. Conclusions

This paper has presented some principles concerning composite decision support based on combining cost-benefit analysis (CBA) with multi-criteria decision analysis (MCDA). Specifically a composite model for assessment (COSIMA) has been presented as a decision support system (DSS). The major focus has been exposing the potential of the COSIMA DSS as a tool for complex assessment problems, which has been assisted by illuminating it with a case example. The following characteristics of the COSIMA DSS can be noted. COSIMA is simple in its design and application compared to earlier attempts to composite analyses (e.g. EUNET(2001) or Tsamboulas (2007)), as the methodology basically just “adds to” and does not hide or change the information given by the CBA. Furthermore, it contains qualities that make it suitable for handling complex assessment problems by incorporation of relevant MCDA-criteria and applications based on weights. In this way the methodology behind COSIMA sets-out guidelines for dealing with the overall feasibility issues of a project appraisal by exploring whether other issues or criteria complementing the CBA can make a project change from being non-feasible to attractive. The methodology has been formulated to deal with the often occurring issue that the CBA result is not sufficient for the actual problem as decision-makers often want additional, systematic examinations that can supplement the CBA. In this respect the COSIMA methodology will be useful and, furthermore, the approach with its new features may be perceived easier accessible by the decision-makers than more complex types of MCDA.

The COSIMA DSS differs from previous attempts on doing composite appraisals in the transport sector in several ways. First of all the COSIMA DSS seeks to ‘translate’ the MCDA results into the same ‘language’ as the CBA results make it possible to produce a total rate of return (TRR), whereas most recent methodologies incorporate the CBA in the MCDA. Obviously, the TRR outcome from the composite expression has no economic argument even though expressed similar to the benefit cost rate. Instead the TRR describes the attractiveness of the alternative seen from both the CBA and MCDA. Thus, the innovative advantage of using the COSIMA approach is that the CBA results are maintained throughout the analysis. Moreover, COSIMA has the advantage that expressing the outcome on a graph as depicted in Figure 3 makes it possible to review the results sensitivity with regard to the weights assigned to the CBA and MCDA respectively.

Overall, it can be concluded that COSIMA contribute in a new way to make decisions more informed. It is moreover seen as a major feature of the modelling approach that the various inputs needed from the decision-makers can help trigger important discussions. This issue has not been discussed thoroughly in this paper, but the outlined decision conference is a method to support and facilitate these discussions amongst decision-makers as described by Phillips (2006) and treated further in Mustajoki et al. (2007). A future research task will thus be to explore the modelling and decision-maker interaction further with the purpose of improving the learning and understanding among the decision-makers about the actual non-standard appraisal task.

References

Banister, D. and Berechman, J. (2000). *Transport Investment and Economic Development*. UCL Press, London, United Kingdom.

Barfod, M.B. (2006). Appraisal of alternatives for a new Hornsherred connection. Master Thesis, Center for Traffic and Transport, Technical University of Denmark.

Barzilai, J., Cook, W.D. and Golany, B. (1987). Consistent Weights for Judgment Matrices of the Relative Importance of Alternatives. *Operations Research Letters* 6 (3), pp. 131-134, Elsevier.

Belton, V. and Stewart, T.J. (2002). *Multi Criteria Decision Analysis: An Integrated Approach*. Kluwer Academic Publishers, London, United Kingdom.

Chen, Y., Kilgour, D.M. and Hipel, K.W. (2008). Screening in multiple criteria decision analysis. *Decision Support Systems* 45, pp. 278-290, Elsevier.

Danish Ministry of Transport (2003). Manual for Socio Economic Appraisal: Applied Methodology in the Transport Sector (in Danish). Danish Ministry of Transport, Copenhagen, Denmark.

EUNET/SASI (2001). Final Report – Executive Summary. 4th RTD Framework Programme of the European Commission.

van Exel, J., Rienstra, S., Gommers, M., Pearman, A. and Tsamboulas, D. (2002). EU involvement in TEN development: network effects and European value added. *Transport Policy* 9, pp. 299-311, Elsevier.

Goodwin, P. and Wright, G. (2009). *Decision Analysis for Management Judgment*. Fourth edition, Wiley, Chichester, United Kingdom.

Hayashi, Y. and Morisugi, H. (2000). International comparison of background concept and methodology of transport project appraisal. *Transport Policy* 7, pp. 73-88, Elsevier.

Hwang, C.L. and Yoon, K. (1995). *Multi Attribute Decision Making – An Introduction*. Sage University Papers, Series/Number 07-104, Sage Publications, pp. 59-66.

Janic, M. (2003). Multi-criteria Evaluation of High-speed Rail, Transrapid Maglev and Air Passenger Transport in Europe. *Transportation Planning and Technology* 26 (6), pp. 491-512, Taylor & Francis.

Jeppesen, S. L. (2010). Sustainable Transport Planning – A Multi-Methodology Approach to Decision Making. Ph.D. thesis, Department of Transport, Technical University of Denmark.

Jiménez, A., Ríos-Insua, S. and Mateos, A. (2003). A decision support system for multiattribute utility evaluation based on imprecise assignments. *Decision Support Systems* 36, pp. 65-79, Elsevier.

Keeney, R.L. and Raiffa, H. (1993). *Decisions with Multiple Objectives – Preferences and Value Tradeoffs*. Cambridge University Press, New York, United States of America.

Larichev, O.I., Kortnev, A.V. and Kochin, D.Y. (2002). Decision support system for classification of a finite set of multicriteria alternatives. *Decision Support Systems* 33, pp. 13-21, Elsevier.

Leleur, S. (2000). *Road Infrastructure Planning: A Decision-Oriented Approach*. Second edition, Polyteknisk Press, Kgs. Lyngby, Denmark.

Leleur, S. (2008). *Systemic Planning*. Second edition, Polyteknisk Press, Kgs. Lyngby, Denmark.

Lootsma, F.A. (1999). *Multi-criteria decision analysis via ratio and difference judgement*. Kluwer Academic Publishers, Dordrecht, Netherlands.

Mackie, P. and Preston, J. (1998). Twenty-one sources of error and bias in transport project appraisal. *Transport Policy* 5, pp. 1-7, Elsevier.

Mustajoki, J., Hämäläinen, R. P. and Sinkko, K. (2007). Interactive computer support in decision conferencing: Two cases on off-site nuclear emergency management. *Decision Support Systems* 42, pp. 2247-2260, Elsevier.

Olson, D.L., Fliedner, G. and Currie, K. (1995). Comparison of the REMBRANDT system with analytic hierarchy process. *European Journal of Operational Research* 82, pp. 522-539, Elsevier.

Phillips, L.D. (1984). A theory of requisite decision models. *Acta Psychologica* 56, pp. 29-48, Elsevier.

Phillips, L.D. (2006). Decision Conferencing. A Working Paper from London School of Economics and Political Science (06.85), Operational Research Group, Department of Management, United Kingdom.

Roberts, R. and Goodwin, P. (2002). Weight Approximations in Multi-attribute Decision Models. *Journal of Multi-Criteria Decision Analysis* 11, pp. 291-303, Wiley.

Saaty, T.L. (1977). A Scaling Method for Priorities in Hierarchical Structures. *Journal of Mathematical Psychology* 15, pp. 234-281, Elsevier.

Saaty, T. L. (2001). *Decision Making for Leaders: The Analytic Hierarchy Process for Decisions in a Complex World*. The Analytic Hierarchy Process Series, Vol. 2, Pittsburgh, United States of America.

Salling, K.B. (2008). Assessment of Transport Projects: Risk Analysis and Decision Support. Ph.D. Thesis, Department of Transport, Technical University of Denmark.

Salling, K.B., Leleur, S. and Jensen, A.V. (2007). Modelling Decision Support and Uncertainty for Large Transport Infrastructure Projects: The CLG-DSS model of the Øresund Fixed Link. *Decision Support Systems* 43, pp. 1539-1547, Elsevier.

Sayers, T.M., Jessop, A.T. and Hills, P.J. (2003) Multi-criteria evaluation of transport options – flexible, transparent and user-friendly. *Transport Policy* 10, pp. 95-105, Elsevier.

Tsamboulas, D.A. and Mikroudis, G.K. (2000). EFECT – evaluation framework of environmental impacts and costs of transport initiatives. *Transportation Research Part D* 5, pp. 283-303, Elsevier.

Tsamboulas, D.A. and Mikroudis, G.K. (2006). TRANS-POL: A mediator between transportation models and decision makers' policies. *Decision Support Systems* 42, pp. 879-897, Elsevier.

Tsamboulas, D.A. (2007). A tool for prioritizing multinational transport infrastructure investments. *Transport Policy* 14, pp. 11-26, Elsevier.

Vickerman, R. (2000). Evaluation methodologies for transport projects in the United Kingdom. *Transport Policy* 7, pp. 7-16, Elsevier.

Vreeker, R., Nijkamp, P. and Welle, C.T. (2002). A multicriteria decision support methodology for evaluating airport expansion plans. *Transportation Research Part D* 7, pp. 27-47, Elsevier.

von Winterfeldt, D. and Edwards, W. (1986). *Decision Analysis and Behavioral Research*. Cambridge University Press, New York, United States of America.

Wright, S., Nelson, J.D., Cooper, J.M. and Murphy, S. (2009). An evaluation of the transport to employment (T2E) scheme in Highland Scotland using social return on investment (SROI). *Journal of Transport Geography* 17, pp. 457-467, Elsevier.

Paper 2

Examination of decision support systems for composite CBA and MCDA assessments of transport infrastructure projects

Michael Bruhn Barfod
Department of Transport, Technical University of Denmark

Anders Vestergaard Jensen
Department of Transport, Technical University of Denmark

Steen Leleur
Department of Transport, Technical University of Denmark

Paper published in Y. Shi et al. (Eds.). *New State of MCDM in the 21st century – Selected Papers of the 20th International Conference on Multiple Criteria Decision Making 2009.* Lecture Notes in Economics and Mathematical Systems, Volume 648, pp. 167-176, Springer

Keywords: Decision support systems, multi-criteria decision analysis, COSIMA, REMBRANDT

Abstract

This paper examines decision support systems (DSS) for composite appraisals of transport infrastructure projects comprising both cost-benefit analysis (CBA) and multi-criteria decision analysis (MCDA). Two DSS, REMBRANDT and COSIMA, are in this context examined and compared using a case study dealing with alternatives for a new high-speed railway line in Sweden. The REMBRANDT system is based on multiplicative value functions and makes use of pair wise comparisons on both attribute and criteria level. The COSIMA system is based on additive value functions and makes use of the REMBRANDT technique using pair wise comparisons on attribute level and swing weights on criteria level. One difference between the two approaches is the focus the COSIMA system puts on combining the CBA and MCDA results influencing, among other things, the way that the final results are expressed. Finally, a recommendation for the use of DSS within transport infrastructure appraisals is set out.

1. Introduction

Two DSS, REMBRANDT (Ratio Estimation in Magnitudes or deci-Bells to Rate Alternatives which are Non-DominaTed) and COSIMA (Composite Model for Assessment), are compared with the purpose of identifying the most appropriate DSS for transport infrastructure assessments including both CBA and MCDA. The first DSS examined, which is widely used and based on acknowledged methods, comprises the REMBRANDT technique (Lootsma, 1992) using pair wise comparisons for rating of the alternatives and determination of the criteria weights. The results of the CBA are in this system compared and included as an additional criterion in the MCDA. Hence, the result of the system is a relative weight-score for each alternative reflecting its performance in the composite appraisal.

The second DSS examined, the so-called COSIMA approach (Leleur, 2000; Salling et al., 2007), provides a framework for adding value functions determined in a MCDA to impacts monetarily assessed in a CBA. The DSS comprises the REMBRANDT technique (Lootsma, 1992) using pair wise comparisons for rating of the alternatives and swing weights (von Winterfeldt and Edwards, 1986) for the determination of criteria weights. However, the COSIMA system does not convert the CBA into an additional MCDA criterion. Instead the value functions computed in the MCDA are added to the CBA results using a balance indicator assigning shadow prices to the MCDA criteria. Subsequently, the resulting total value is divided by the investment costs. Hence, the result is a total rate for each alternative reflecting its attractiveness in the appraisal as a function of the weight-set between the CBA and MCDA.

The input for the two DSS examined was generated using a case study. For this purpose a decision conference (Phillips, 2006) was set up where various stakeholders and decision-makers under the guidance of a facilitator were producing input to the DSS in form of their preferences.

The purpose of the examination and comparison of the two DSS is to determine which is the most appropriate for conducting composite appraisals of transport infrastructure projects. The REMBRANDT system provides a conventional widely used and theoretical well founded framework, while the COSIMA framework is more recent and founded on a somewhat different set of axioms. However, the two systems provide the decision-makers with the same type of result, only expressed differently. The question treated in this context is hence if the COSIMA system provides the decision-makers with some information that the REMBRANDT system does not provide and vice versa.

Finally, conclusions are drawn including a recommendation based on the case study for the most appropriate system for conducting composite appraisals of transport infrastructure projects, and research questions defining future work in the context of composite DSS and their use in decision making processes are set out.

2. Value measurement

Value measurement theory is introduced in the following in order to underpin the theory behind the two DSS, which makes use of different types of value function methods and scales.

Value function methods basically produce the assessments of the performance of alternatives against individual criteria. This together with inter-criteria information reflecting the relative importance of the different criteria, w_j, makes it possible to give an overall evaluation of each alternative indicative of the decision-makers preferences. The simplest and most widely used form of value function method, which is used by the COSIMA system, is the additive model (Belton and Stewart, 2002):

$$V(a) = \sum_{j=1}^{m} w_j v_j(a) \tag{1}$$

Considerably more complicated in appearance, but as easy to use, is the multiplicative model which the REMBRANDT system makes use of (Ibid):

$$V(a) = \prod_{j=1}^{m} \left[v_j(a)\right]^{w_j} \tag{2}$$

In its analytical expansion the multiplicative model seems prohibitive compared to the additive model. However, it requires only the addition of a single parameter (w), which defines all interaction terms. Therefore, the type of interaction it models is rather constrained (von Winterfeldt and Edwards, 1986). Moreover, additive aggregation is the form that is most easily explained and understood by decision-makers from a wide variety of backgrounds, while not placing any substantially greater restrictions on the preference structures than more complicated aggregation formulae (Belton and Stewart, 2002).

If the criteria are structured as a value tree (for details see for instance Goodwin and Wright (2004)) then the alternatives must be scored against each of the bottom level criteria. These values need to be assessed on an interval measurement scale, i.e. a scale on which the difference between the points is the key factor. A ratio of values will only be meaningful if the zero point on the scale is absolutely and unambiguously defined. Thus to construct a scale it is necessary to define two reference points and to assign numerical values to these points. The minimum and maximum points on the scale can be defined in a number of ways (here 0 and 100). However, it is very important to distinguish between a local and a global scale.

A local scale is defined by the set of alternatives that are under consideration. The alternative which performs best within a specific criterion is for instance assigned the score 100 and the alternative which performs least well is assigned the score 0. All other alternatives will then receive intermediate scores which reflect their performance relative to the end points. The use of local scales allows a relative fast assessment of values and it can be very useful for preliminary "roughing out" of a problem, or if operating under time constraints. However, some issues in the context of local scales will be discussed later.

A global scale is defined by reference to a broader set of possibilities. The end points may be defined by the ideal and the worst possible performance within the particular criterion (extreme endpoints), or by the best and worst performance that can realistically occur. The definition of a global scale requires more preparatory work than a local scale. However, it has the advantages that it is more general than a local scale and that it can be defined before consideration of specific alternatives. This also means that it is possible to define criteria weights before consideration of alternatives.

The important point is that subsequent analysis, including the assessment of the weights (w_j), must be consistent with the chosen scaling. Once the reference points of the scale have been determined consideration must be given to how other scores are to be assessed. In this paper it has been chosen to use direct rating of the alternatives using pair wise comparisons. The pair wise comparisons have shown to be a strong decision aid when making decisions in groups, and hence are appropriate for use at the decision conference in the treated case study.

3. The case study

The case study examined concerns an assessment of alternatives for a new high-speed railway line in Sweden named Ostlänken. More specific the case study considers a section between Bäckeby and Norrköping about 100 kilometres south of Stockholm. The work described in this paper was done as a part of a research project granted by the Swedish Research Council – VINNOVA.

Four alternatives describing different alignments were to be compared. These are in the following referred to as alternative R, BS, BL and G. A conventional cost-benefit analysis (CBA) in accordance with a national socio-economic manual was carried out at a preliminary stage. The calculations were carried out according to Swedish standards (Hiselius et al., 2009), and the outcome expressed as benefit-cost rates (BCR) are shown in Table 1.

Table 1. CBA information for the four alternatives

	R	BS	BL	G	Method	Unit
Costs	1509	1774	2033	2167	CBA	M SEK
Benefits	3018	3138	3138	3140	CBA	M SEK
BCR	2.00	1.77	1.54	1.45	CBA	

The results clearly indicate that alternative R is the economically most feasible project. However, all four alternatives were beneficial seen from a CBA based point of view. The decision-makers decided to complement the assessment with some strategic (non-monetary) impacts as well, as they felt that the CBA did not cover all aspects of the decision problem. Hence, there was a need for a more comprehensive assessment. Eight different criteria were in this context defined describing what was lacking in the CBA. Effort was made to avoid double counting; however, the decision-makers were aware of the risk of this. The eight complementing criteria will in the following be referred to as C1 to C8.

Using a decision conference as proposed by Phillips (2006) the input needed for the two DSS were produced. The participants were experts with extensive knowledge within the respective criteria and each assessment made was documented in a protocol (developed for the purpose). This procedure makes it possible for decision-makers as well as the general public to review each judgment made; thereby the decision process becomes more transparent.

The following two sections describe the two DSS applied to the case study and their results are presented.

3.1 The REMBRANDT system

A systematic pair wise comparison approach is one of the cornerstones of the REMBRANDT system by Lootsma (1992). REMBRANDT makes use of a procedure for direct rating which requires the decision-makers to consider all possible pairs of alternatives with respect to each criterion in turn, in order to determine which of the pair is preferred and to specify the strength of preference according to a semantic scale (or the associated numeric 0-8 scale). The approach is a further development of the Analytic Hierarchy Process (AHP) by Saaty (1977) and it proposes to overcome three issues regarding the theory behind AHP.

First, the direct rating in REMBRANDT is on a logarithmic scale (Lootsma, 1988) which replaces Saaty's 1 – 9 fundamental scale. Second, the eigenvector method originally used in AHP is replaced by the geometric mean, which avoids potential rank reversal (Barzilai et al., 1987). Third, the aggregation of scores by arithmetic mean is replaced by the product of alternative relative scores weighted by the power of weights obtained from analysis of hierarchical elements above the alternatives (Olson, 1996). The differences between AHP and REMBRANDT have been treated very thoroughly by Olson et al. (1995) and will for this reason not be treated any further in this paper.

In the REMBRANDT system the ratio value r_{jk} on the geometric scale is expressed as an exponential function of the difference between the echelons of value on the geometric scale δ_{jk}, as well as a scale parameter γ. Lootsma considers two alternative scale parameters γ to express preferences. For calculating the weight of criteria, $\gamma = \ln\sqrt{2} \approx 0.347$ is used. For calculating the weight of alternatives on each criterion, $\gamma = \ln 2 \approx 0.693$ is used.

The participants at the decision conference were using the REMBRANDT technique based on pair wise comparisons of the four alternatives within all eight criteria (C1 to C8). Moreover an additional economic efficiency criterion (C9), was added as the alternatives were also compared on the basis of their BCR. Then scores were calculated within all nine criteria (C1 to C9) using logarithmic regression and the geometric mean method. Finally, criteria weights were determined using pair wise comparisons as well, and aggregated values for the alternatives were calculated. The participants decided to test 3 different weight-sets at 0.85, 0.70 and 0.55 (derived by the pair wise comparisons) where the relative weight of the BCR-criterion was changed: This was done to test the robustness of the result, see Table 2.

Table 2. Results from REMBRANDT

Total BCR weight – relative value	0.85	0.70	0.55
R	0.316	0.246	0.184
BS	**0.319**	**0.354**	**0.380**
BL	0.214	0.267	0.322
G	0.151	0.133	0.114

The results in Table 2 depict BS as the most attractive alternative within the three weight-sets as this alternative obtain the highest relative values. However, if testing other weight-sets the resulting ranking could be different. The participants at the decision conference were confident that the total BCR-weight should be somewhere within the investigated interval. Hence, BS proved to be a robust choice when combining the CBA and MCDA results in the REMBRANDT DSS.

3.2 The COSIMA system

The COSIMA system is based on adding non-monetary MCDA-criteria to the monetary CBA-impacts. The model uses the argument that the MCDA-criteria are additive to the CBA-impacts if proper value-functions for the MCDA-criteria are computed and assigned with shadow prices describing each criterion's importance.

COSIMA is developed with the purpose of handling a situation where criteria (strategic impacts) cannot be monetised in a comprehensive and transparent manner. This aspect is in COSIMA ensured through the determination of appropriate weights for the MCDA-criteria and appropriate value function scores for the alternatives. The COSIMA system has proven its worth for providing efficient decision support in various types of infrastructure projects; for example the Copenhagen-Ringsted railway line (Salling et al., 2008) and the Øresund fixed link (Salling et al., 2007). Unlike the REMBRANDT system the COSIMA system does not include the BCRs in the MCDA module. Instead, COSIMA "translates" the MCDA-part into a

CBA-like ‘language’ keeping the economic information provided by the BCRs intact at all times. For readers interested in the specific calculations – not accounted for here – see Leleur (2008) or Hiselius et al. (2009).

COSIMA features several possible options for assigning scores to the alternatives and weights to the criteria. More or less any MCDA-technique can be applied to the system. For the case study it was, however, decided to apply the REMBRANDT technique for the scoring of alternatives (similar to section 3.1) and the swing weight technique (von Winterfeldt and Edwards, 1986) for the determination of criteria weights.

The resulting scores from the pair wise comparisons within the eight MCDA-criteria were consequently transformed into value function scores applying the previously mentioned local scale using a linear assumption, see Table 3.

Table 3. Value function scores derived using the REMBRANDT technique

	C1	C2	C3	C4	C5	C6	C7	C8
R	1	0	35	0	0	11	0	39
BS	100	18	35	100	100	5	100	0
BL	70	25	100	51	100	0	100	0
G	0	100	0	15	55	100	2	100

The swing weight technique was used for the determination of criteria weights. The participants at the decision conference were asked to rank the criteria in order of importance and subsequently assess the value of the swing from best to worst performance within each criterion in turn compared to the swing from best to worst within the highest ranked criterion. The weights are shown in Table 4.

Table 4. Swing weights for the eight criteria

Criteria	Swing weights	Normalised weights
C1	90	0.19
C2	80	0.17
C3	100	0.21
C4	60	0.12
C5	70	0.14
C6	20	0.04
C7	60	0.12
C8	5	0.01

The participants had already decided to test 3 different weight-sets where the relative weight of the BCR compared to the MCDA-part was the varying parameter (see section 3.1). The COSIMA DSS was consequently calibrated in accordance with the weights in Table 4 and the balance (trade-off) between MCDA and CBA. Summarising the calculations and the process concerning the case it can be noted that the analysis is based on the use of cost-benefit analysis (CBA) using a national manual and multi-criteria analysis (MCDA) using a decision conference. The CBA produces results that can be measured in monetary units – here million Swedish Kroner (M SEK). The MCDA on the other hand produces results that by comparison with the CBA results can be calibrated to 'assessment M SEK'. In order to obtain a total value for an examined alternative M SEK and 'assessment M SEK' is added. This mix between M SEK and the fictitious 'assessment M SEK' is expressed by the unit 'attractiveness M SEK'. The result can also be presented as a total rate of return (TRR), where the result in 'attractiveness M SEK' is divided by the investment costs, see Table 5. In this context it should be noted that the process based on input (scoring of alternatives, determination of criteria-weights and balancing the CBA and MCDA) makes it possible to provide a transparent evaluation process involving the decision-makers. Thus, the result of the COSIMA analysis is that it by use of stakeholder and decision-maker involvement is possible to add a value to the MCDA-criteria which can be compared to the monetary CBA values. Thereby, the result scheme in Table 5 indicates the 'gain' by choosing an alternative that performs well within the MCDA instead of the alternative that performs the worst. Thus, seen from a strategic point of view the decision-makers in the current case study would gain most from the investment by choosing the BS alternative.

Table 5. Result scheme from the COSIMA analysis with the CBA weight set to be 0.70

	R	BS	BL	G	Method	Unit
Costs	1509	1774	2033	2167	CBA	M SEK
Benefits	3018	3138	3138	3140	CBA	M SEK
B/C rate	2.00	1.77	1.54	1.45	CBA	
C1	7	555	387	0	MCDA	'Assessment M SEK'
C2	0	87	123	493	MCDA	'Assessment M SEK'
C3	215	215	616	0	MCDA	'Assessment M SEK'
C4	0	370	188	54	MCDA	'Assessment M SEK'
C5	0	431	431	236	MCDA	'Assessment M SEK'
C6	14	6	0	123	MCDA	'Assessment M SEK'
C7	0	370	370	8	MCDA	'Assessment M SEK'
C8	12	0	0	31	MCDA	'Assessment M SEK'
Total MCDA	248	2033	2114	945	MCDA	'Assessment M SEK'
Total value	3266	5171	5252	4058	CBA+MCDA	'Attractiveness M SEK'
Total rate	2.16	2.91	2.58	1.88		

3.3 Comparison of the two DSS

Two DSS have in this paper been applied to the same case study revealing the same results. In Table 6 the rankings of the alternatives at the different weight-sets are depicted for both DSS. This clearly shows that the two DSS also provide the same results on second and third level in the rankings.

Table 6. Rankings of the alternatives at different weight-sets

CBA weight	**0.85**		**0.70**		**0.55**	
	REMBRANDT	COSIMA	REMBRANDT	COSIMA	REMBRANDT	COSIMA
R	2	2	3	3	3	3
BS	1	1	1	1	1	1
BL	3	3	2	2	2	2
G	4	4	4	4	4	4

The difference on the results of the two DSS, however, consists in the way they are expressed. The REMBRANDT DSS provides the decision-makers with weight-scores expressing the alternatives relative performance against each other. The COSIMA DSS on the other hand provides the decision-makers with a somewhat more informed result. The total rate (TRR) from COSIMA features both the CBA result and the MCDA result expressed in one single rate. The REMBRANDT system is a theoretical well founded system which have been applied to various decision problems, see e.g. (Van den Honert and Lootsma, 2000), and on which other systems can be measured. Given that the COSIMA DSS provides the same result (in this case) as REMBRANDT the DSS seems most appropriate for use within transport infrastructure planning as the results provide the decision-makers with two-way information containing both an economic argument and a strategic argument. This makes the COSIMA results more useful especial when the results need to be transparent and defendable to the public.

However, the use of local scales within the COSIMA system arises some issues. Firstly, projects assessed using this scale cannot easily be compared with other projects as the scale presupposes a closed system. This way the endpoints within the assessment define the scale, in contradiction to a global scale which considers the extreme endpoints in defining the scale. If projects are not assessed using the same scale it is obvious that they cannot be compared to each other. Secondly, the segregation between the attributes within some criteria can in some cases almost be negligible. However, the use of a local scale will anyway imply large differences between these attributes. E.g. in a case where $a_1 = 0.33$, $a_2 = 0.33$ and $a_3 = 0.34$ the value function scores will be 0 for a_1 and a_2, but a_3 will obtain the score 100. This does not seem reasonable as the alternatives perform almost identical in the assessment of the criterion. A means to overcome this problem is to perform a check of all the criteria in the assessment using the swing weight method. If this check reveals some criteria where the swing from the best to the worst performing alternative almost does not exist (a lower boundary can be set by the decision-makers), the criteria should be removed from the analysis as they do not contribute to the segregation between the alternatives and therefore are without significance for the appraisal task in hand.

Furthermore, it is most likely that the weightings of the criteria will diverge dependent of the decision-maker or stakeholder that makes the judgments. Thus, this level of the assessment can be seen as more subjective than the attribute level where the issues are broken down into simple objective judgments within each criterion. Therefore, it can be beneficial to evaluate different stakeholder's preferences in order to obtain a broader perspective for the final decision making.

4. Conclusion and perspective

This paper has compared the COSIMA and REMBRANDT DSS with the purpose of determining which of the techniques are most appropriate for decision making within transport infrastructure projects. Both systems seem well suited for decision making in groups, thus the differences between the systems relate to their procedural operations. COSIMA and REMBRANDT are both considered to be effective tools for aiding decision problems (especially in groups) faced with multiple criteria. However, for the use within transport planning the COSIMA DSS seems most appropriate. The COSIMA DSS provides the decision-makers with a more informed result as the TRR express both the feasibility using the BCR and the added value of the assessed MCDA-criteria.

The COSIMA DSS, moreover, provides the opportunity for using other MCDA techniques depending on the character of the assessment task. A future research task is in this respect to outline a framework of MCDA techniques that are appropriate to apply to COSIMA at different levels of complexity and at the same time be used by a decision conference.

References

Barzilai, J., Cook, W. and Golany, B. (1987). Consistent weights for judgement matrices of the relative importance for alternatives. *Operations Research Letters* 6:3, pp. 131-134, Elsevier.

Belton, V. and Stewart, T.J. (2002). *Multi Criteria Decision Analysis – An Integrated Approach*. Kluwer Academic Publishers, London, United Kingdom.

Goodwin, P. and Wright, G. (2004). *Decision Analysis for Management Judgment*. Third Edition, Wiley, Chichester, United Kingdom.

Hiselius, L., Barfod, M.B., Leleur, S., Jeppesen, S.L., Jensen, A.V. and Hjalte, K. (2009). Helhetsorienterad Utvärdering av Kollektivtrafikåtgärder. University of Lund, Sweden.

Leleur, S. (2000). *Road Infrastructure Planning – A Decision Oriented Approach*. Second edition, Polyteknisk Press, Kgs. Lyngby, Denmark.

Leleur, S. (2008). *Systemic Planning*. Second edition, Polyteknisk Press, Kgs. Lyngby, Denmark.

Lootsma, F.A. (1988). Numerical scaling of human judgement in pairwise-comparison methods for fuzzy multi-criteria decision analysis. In: Mitra, G. (ed.). *Mathematical Models for Decision Support* (pp. 57-88). Springer-Verlag.

Lootsma, F.A. (1992). The REMBRANDT system for multi-criteria decision analysis via pair wise comparisons or direct rating. Report 92-05, Faculteit der Technische Wiskunde en Informatica, Delft University of Technology, Netherlands.

Olson, D.L. (1996). *Decision Aids for Selection Problems*. Springer series in Operations Research, Springer-Verlag, New York, United States of America.

Olson, D.L., Fliedner, G. and Currie, K. (1995). Comparison of the REMBRANDT system with analytic hierarchy process. *European Journal of Operational Research* 82, pp. 522-539, Elsevier.

Phillips, L. D. (2006). Decision Conferencing. A working paper LSEOR 06.85, London, United Kingdom.

Saaty, T. L. (1977). Scenarios and priorities in transport planning: Application to the Sudan. *Transport Research* 11, pp. 343-350, Pergamon Press.

Salling, K.B., Leleur, S. and Jensen, A.V. (2007): Modelling decision support and uncertainty for large infrastructure projects: The CLG-DSS Model of the Øresund fixed link. *Decision Support Systems* 43, pp. 1539-1547, Elsevier.

Salling, K.B., Landex, A. and Barfod, M.B. (2008). Appraisal of the railway line between Copenhagen and Ringsted by the use of a decision support model named COSIMA-DSS. Accepted for publication in *Journal of Advanced Transportation*.

van den Honert, R.C. and Lootsma, F.A. (2000). Assessing the quality of negotiated proposals using the REMBRANDT system. *European Journal of Operations Research* 120, pp. 162-173, Elsevier.

von Winterfeldt, D., and Edwards, W. (1986). Decision Analysis and Behavioral Research. Cambridge University Press, New York, United States of America.

Paper 3

Customised DSS and decision conferences

Michael Bruhn Barfod
Department of Transport, Technical University of Denmark

Steen Leleur
Department of Transport, Technical University of Denmark

Paper published in the Proceedings of the 13th Euro Working Group on Transportation Meeting 2009 (EWGT2009), University of Padova, Italy.

Keywords: Decision conference, decision support systems, multi-criteria decision analysis, REMBRANDT, swing weights

Abstract

This paper presents and exemplifies a combination of techniques for deriving and modelling decision-maker and/or stakeholder preferences using a decision conference process. The applied techniques are used for the development of customised decision support systems (C-DSS) which can be used for appraisals of large transport infrastructure projects. The paper exemplifies how the process at a decision conference can be effectively supported by a DSS customised using appropriate techniques for the specific task in hand. In this respect a conventional cost-benefit analysis (CBA) is combined with a multi-criteria decision analysis (MCDA) featuring the REMBRANDT and the swing weights techniques. The approach is presented based on a case study, which concerns the interaction between stakeholders and decision-makers at a decision conference which was set up for the appraisal of proposals for the alignment of a high-speed railway line in Sweden.

1. Introduction

The concept of a decision conference is introduced into transport infrastructure planning as the decisions to be made in this context often are of a very complex character as various stakeholders and authorities seem to have great leverage in the debate concerning these types of projects and hence also in the final decision. Thus a need has arisen for a structured decision process which can take all aspects into account and at the same time be transparent both to the participants and the public. This paper provides a proposal for how such a process can be outlined both with regard to the decision-maker interaction and the techniques used in the underlying customised decision support system (C-DSS).

The C-DSS introduced is customised to the specific assessment task using techniques that reflect the current needs and composition of the decision-makers and/or stakeholders participating in the decision process. This means that the C-DSS is constructed through an interactive and consultative process between problem owners and specialists (Phillips, 1984). The decision-maker interaction with the C-DSS is proposed to take place using a decision conference as described by Phillips and Bana e Costa (2005) and Phillips (2006), where key players are brought together under the guidance of a facilitator with the purpose of discussing and assessing the relevant issues.

The paper is composed as follows. The need for using multi-criteria decision analysis (MCDA) within transport planning is introduced together with the methodological background of the C-DSS. Next, the concept of decision conferences is described introducing the five step procedure that is proposed to structure a decision making process concerning traffic infrastructure projects. A case study dealing with a large infrastructure project in Sweden is following used to demonstrate the procedure and the techniques being used by the C-DSS. Finally, the advantages and disadvantages of preference modelling within composite project appraisals are discussed, a conclusion is drawn and a perspective for the future work is given.

2. Methodology

The traditions for appraisals of transport infrastructure projects differ from country to country; however a conventional cost-benefit analysis (CBA) is to some extent conducted in most countries. The CBA provides the decision-makers with a monetary assessment of the profitability of the project alternatives. However, it is widely accepted that the decision making regarding infrastructure projects (and many other types of decision problems) often are influenced by some more strategic impacts as well, which have the possibility of improving the basis for decision. In order to assess these impacts, which often cannot be monetised, the concept of multi-criteria decision analysis (MCDA) is introduced. In some countries, e.g. France and the Netherlands (Leleur, 2000), MCDA is widely applied, while in other countries with long traditions for CBA, e.g. in Sweden and Denmark, the concept is not yet fully accepted as a valid decision aid.

A comprehensive type of MCDA is introduced in order to attain an overall result for the assessment task comprising all relevant impacts. The MCDA is based on the theoretically well founded additive value function (von Winterfeldt and Edwards, 1986; Keeney and Raiffa, 1993), which using the case study is explored by a weight assignment procedure for determining the importance of the criteria. The conventional CBA results are included in the MCDA by assigning value function scores to the B/C rates as described in e.g. EUNET (2001). The REMBRANDT technique (Lootsma, 1992; Olson et al., 1995) based on a full set of pair wise comparisons is used for the determination of value-function scores (Belton and Stewart, 2002) for the alternatives within the strategic non-monetised criteria, and swing weights (von Winterfeldt and Edwards, 1986) are applied in order to explore the decision-maker interaction when assigning weights to the criteria. Hence, the composite assessment of the CBA and MCDA seeks to clarify, how the final decision making partly depends on which strategic preferences the decision-makers choose as foundation for assessing the issue, and partly on how much weighting the MCDA-criteria is assigned compared to the CBA-impacts. The components leading to the comprehensive assessment within the C-DSS are shown in Figure 1

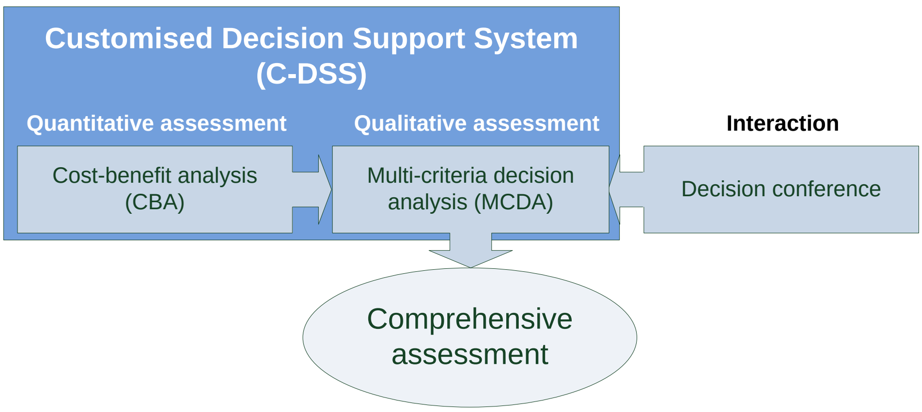


Figure 1. The components leading to a comprehensive assessment

Figure 1 depicts that the quantitative CBA is embodied in the qualitative MCDA within the C-DSS. Moreover, to support the techniques and lead to a comprehensive assessment, the process of interacting with the decision-makers/stakeholders is considered. This can be referred to as the decision conference component of the assessment. It is described how decision-makers/stakeholders can be involved in the decision process and thereby express their preferences in designing the specific C-DSS, which by intervention of analysts, is set up to address the specific task. Thus, a decision conference joins decision analysis, group processes and information technology in an intensive session where people on various levels that are involved in the decision process ideally are present (Goodwin and Wright, 2004). Basically, a decision conference is an approach that makes it possible for a group of stakeholders and decision-makers representing very different viewpoints to work together in a way so efficiently that they can create a common vision-based decision. The decision conference can for instance take place by a group of decision-makers are being placed around a table with the purpose of discussing the issue. The conference is controlled by a facilitator which organises and facilitates the interplay and knowledge sharing in the group. Moreover, the facilitator is supported by a decision analyst who uses interactive decision support technology (the C-DSS) to model the issues and viewpoints which appear during the process (Phillips, 2006). A decision conference places some high demands on the shoulders of the facilitator. It is the facilitator´s responsibility that everything runs smoothly, that all discussion in the group is constructive and that everybody's opinions are heard. Moreover, the facilitator has to give input to the discussion if it stalls or interfere if one or more of the participants are acting against the rules set up for the conference.

The decision conference helps to conduct the MCDA according to the preferences of the participants in a comprehensive and transparent way. For this reason the process is built up around five simple steps for the participants to consider. Using these five steps it is illustrated how the purpose of the decision conference – creating transparent input to the C-DSS – is obtained.

2.1 The five-step process

The five steps have been formulated in order to be useful in motivating the decision-makers and stakeholders to produce the input needed for the composite assessment in the C-DSS. The steps are as depicted in Figure 2, where the arrows indicate the processes which ensure that the issue and the C-DSS have been completely understood and is treated thoroughly. The arrows pointing back from step 5 indicate that it is possible to go back in the process and redo the assessments made in step 3 and 4 if shared understanding has not been achieved. The steps are described below.

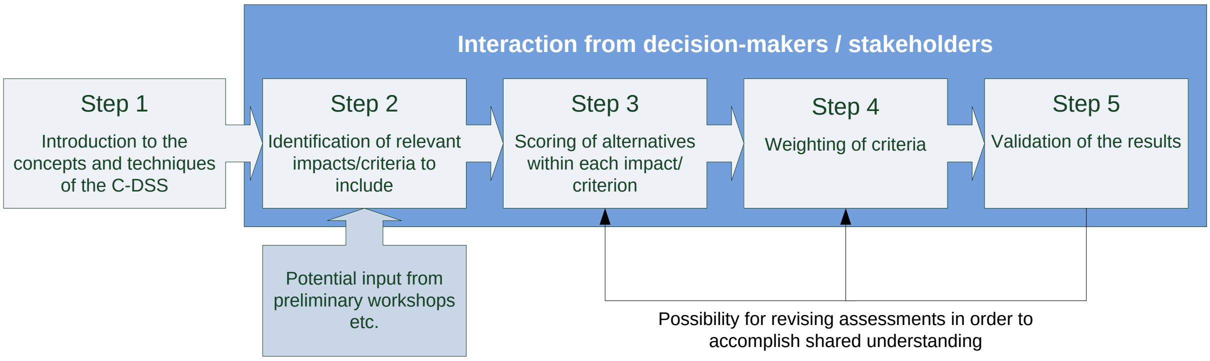


Figure 2. The process at the decision conference comprising five steps

First of all it is very important for the quality of the decision conference that the facilitator starts by introducing the concepts and methods being used in simple terms. If the introduction becomes too technical or theoretical the participants tend to be confused, but a short practical introduction will help them to understand how their inputs are being processed. This contributes to making the decision-makers more comfortable with the later decisions when they know the basic characteristics of the C-DSS. Ideally, the C-DSS is built up in such an intuitive and easily accessible way that the participants do not need a thorough knowledge of the many theories and techniques applied.

The second step features the identification of relevant impacts for the assessment. In this respect it can be very useful to conduct workshops already in the initial planning phase, where issues regarding the project initiative can be discussed and criteria with influence on the decision making can be developed. Moreover, such workshops can help in reducing a high number of project alternatives and instead focusing in a more detailed way on a few.

The types of criteria to develop in this phase highly depend on whether it has already been decided to go forward with the project initiative. If so the decision conference will have the purpose of making an informed choice between alternatives, and in consequence of this the criteria to be included should contribute to the segregation of the alternatives (if all alternatives perform equally under the criterion it should be omitted from the assessment). If, on the other hand, the matter in question is whether to carry forward a project initiative or not (a go/no-go decision) then the criteria to be included will often be of a more strategic economic type. The initial phases might create a lot of different criteria regardless of which of the two before mentioned appraisal tasks that are to be dealt with. Hence, it is up to the participants at the decision conference to structure and reduce the criteria into a number of relevant criteria which all contribute to the differentiation between the alternatives.

Once all relevant criteria have been defined the third step comprises the scoring of the alternatives. Dependent on the level of knowledge about the alternatives and the criterion assessed different techniques can be applied in order to elicit scores. E.g. a criterion like ‘roads distance to natural habitats’ can be measured in meters and a value function score can be assigned directly in accordance with this information, while a criterion like ‘roads visual impact on the landscape’ cannot be measured but scores can still be effectively elicited using pair wise comparisons. A scoring scale should, however, be defined before introducing the step.

While the third step can be seen as something close to being objective – as the criteria should be divided into sub-criteria until ideally no matters of dispute can exist on the alternatives' performance – the fourth step introduces the most subjective part of the appraisal: the weighting of the criteria. The task in this step – to make the participants agree upon a weight set – is considered to be very difficult or even impossible as very opposite world views (which may be present at the decision conference) will create very different weight sets. Instead of trying to make the participants agree, it can be useful to examine the different weight sets provided by each participant individually. These can either point out the same project alternative as being the most attractive, which of course will be the ideal result, or they can point at different project alternatives. If the latter is the case further discussion in the fifth step might lead to common understanding, or the decision-makers might end up with choosing their own favourite alternative. However, the decision-makers will no matter what make their choice based on a broader basis of knowledge after the decision conference, as they are aware of the other stakeholder group's viewpoints and can take these into account.

3. The case study

The case study considered deals with alternatives for a new high-speed railway line named Ostlänken in Sweden. Specifically, the case study considers a section between the two cities Bäckeby and Norrköping about 100 kilometres south of Stockholm. The decision conference described in this paper was carried out as a part of a research project granted by the Swedish Research Council (VINNOVA), see Hiselius et al. (2009) for further details about the project.

In the case study four alternatives describing different alignments for the railway line were to be compared in order to find the most appropriate solution. The alternatives were: R (red alignment), BS (blue alignment with a short tunnel under recreational areas), BL (blue alignment with a long tunnel under recreational areas) and G (green alignment). A decision conference supposed to deal with the issues of the decision problem was scheduled to take place in the winter of 2009. This decision conference was attended by key players and stakeholders all with an extensive knowledge about the project. The decision conference was because of the preparatory work made available from an initial stage (identification of criteria) set to last only half a day. This short type of a decision conference is a possibility if time resources are scarce and participants are experts on their field (Jeppesen, 2009).

3.1 Step one – Introduction

At the beginning of the decision conference the facilitator introduced the participants to the programme of the day stating when interaction was needed from the participants and what the expectations to them were. Moreover, the decision analyst went through the concept and techniques of the underlying C-DSS in an easily accessible manner in order to prevent unnecessary confusion when conducting the assessment later in the process.

Prior to the decision conference a conventional cost-benefit analysis (CBA) was carried out in accordance with the Swedish standard for socio-economic appraisals. The CBA comprised an assessment of the following principal impacts which were estimated in million Swedish Kroner – m SEK (Hiselius et al., 2009):

- Construction, reinvestment and maintenance costs
- Ticket revenues for the operators
- Travel time savings for the users
- Pollution and noise

The outcomes expressed as benefit-cost rates (BCR), see Table 1, were introduced to the participants.

Table 1. Benefit-cost rates for the four alternatives in the appraisal

	R	BS	BL	G	Method	Unit
Costs	1509	1774	2033	2167	CBA	m SEK
Benefits	3018	3138	3138	3140	CBA	m SEK
BCR	2.00	1.77	1.54	1.45	CBA	

The results clearly indicated that alternative R was the economically most profitable project. However, all four alternatives were beneficial seen from this CBA-based point of view. The participants were eager to complement the appraisal with strategic impacts as well, as they felt that the CBA did not cover all aspects of the decision problem. Hence, there was a need for a more comprehensive assessment, where strategic impacts were to be explored by the decision conference.

3.2 Step two – Identification of criteria

As indicated in Figure 2 in section 2.1 preliminary workshops (or other idea generating sessions) might create potential input to the decision conference's second step. For the considered case study an initial Environmental Impact Assessment (EIA) study had already been conducted taking eight strategic issues into account. The participants at the decision conference decided to consider these eight criteria as they were well developed and a lot of resources had been used to describe them in a high level of detail. The criteria (strategic impacts) that were included are:

- C1: City and scenery impression
- C2: Cultural environment
- C3: Natural environment
- C4: Health
- C5: Natural resources
- C6: Risk and safety
- C7: Recreation and outdoor life
- C8: Construction time

3.3 Step three – Scoring of alternatives

The assessment (scoring) of the alternatives were carried out using pair wise comparisons within each criterion. The pair wise comparisons were carried out using the semantic REMBRANDT scale going from indifference to very strong preference for one alternative over another (the associated numerical scale was used as input to the C-DSS). Arguments stating the preference for each of the comparisons were documented in an assessment protocol (developed for the purpose) by a small text describing the two alternatives performance with regard to the criterion. This procedure makes it possible for decision-makers and/or stakeholders as well as "normal" people to review the foundation for each comparison thereby making the decision process more transparent.

The results of the CBA indicated in Table 1 were used for creating an additional criterion, C9, which was based on the BCRs and further processed as one of the eight other criteria. The pair wise comparisons were transformed using the principles of the REMBRANDT technique (Olson et al., 1995) and scores for each alternative were calculated. These scores were then transformed into value-function scores using a linear assumption on a local scale; the score 100 describing the best performing alternative and the score 0 describing the worst performing alternative. As the case study considered can be seen as a so-called closed system – the results are only used to find the "best" solution for one project, the results are not to be used for comparisons with other projects – it seemed reasonable to use a local scale instead of the very time consuming set up of a global scale. The use of local versus global scales is treated in further details in Belton and Stewart (2002).

Table 2 depicts how the numerical values of the pair wise comparisons are transformed using the REMBRANDT scale parameter for alternatives, $\gamma = \ln 2 \approx 0.693$ (Olson et al., 1995), scores are calculated as the geometric mean of the transformed values, and finally the scores are transformed into value function scores.

Table 2. Procession of the pair wise comparisons into value function scores (example for C1)

	R	BS	BL	G	(Transformed using γ = ln2)				Score	Normalised	VF-score
R	0	-6	-4	2	1	0.01563	0.0625	4	0.25	0.02	1
BS	6	0	0	4	64	1	1	16	5.656854	0.56	100
BL	4	0	0	4	16	1	1	16	4	0.40	70
G	-2	-4	-4	0	0.25	0.0625	0.0625	1	0.176777	0.02	0

The same procedure was carried out for all nine criteria, which means the participants conducted 6 x 9 = 54 pair wise comparisons in all. An overview of the value function scores for the alternatives within all nine criteria is shown in Table 3 (C9 is the added criterion based on the BCRs).

Table 3. Value function scores derived using the REMBRANDT technique

	C1	C2	C3	C4	C5	C6	C7	C8	C9
R	1	0	35	0	0	11	0	39	100
BS	100	18	35	100	100	5	100	0	58
BL	70	25	100	51	100	0	100	0	17
G	0	100	0	15	55	100	2	100	0

3.4 Step four – Weighting of criteria

The next task for the participants at the decision conference was to assign weights to the criteria. As the participants were all experts with an extensive knowledge about the criteria it was chosen to use the swing weight technique (Von Winterfeldt and Edwards, 1986), which enables the participants to determine the importance of each criterion very accurately. This part of the assessment – the weighting of the criteria – will most likely diverge dependent on who makes the judgments and their worldview (e.g. the decision-makers or different stakeholders).

Thus, the criteria-level of the assessment can be seen as more subjective than the attribute-level (scoring of alternatives) where the issues are broken down into simple objective judgments under each criterion. Therefore, it can be beneficial in certain cases to evaluate different stakeholders' preferences as, when multiple stakeholders are involved in a decision, there are multiple value structures (weight-sets for the criteria), one for each stakeholder. These are relevant and valuable information can be achieved by modelling each value-structure (Keeney, 1992). However, in the current case study it was chosen only to work with the criteria weights specified by the group decision.

Using the swing weight method the "swing" which is normally considered goes from worst value to best value on each criterion. The participants were asked to consider all the eight original MCDA-criteria simultaneously and to assess which swing gives the highest increase in overall value; this criterion obtains the highest weight. The process was repeated on the remaining set of criteria until the order of benefit resulting from a swing from worst to best on each criterion had been determined, thereby defining a ranking of the criteria weights. To assign values to the weights the participants then had to assess the relative value of the swings. E.g. if a swing from worst to best on the highest ranked criterion is assigned a value of 100 (C3 in Table 4), what is the relative value of a swing from worst to best on the second ranked criterion (C1 in Table 4)? The swing weights derived by the participants are shown in Table 4.

Table 4. Swing weights for criterion C1 – C8

	Swing weights
C3: Natural environment	100
C1: City and scenery impression	90
C2: Cultural environment	80
C5: Natural resources	70
C4: Health	60
C7: Recreation and outdoor life	60
C6: Risk and safety	20
C8: Construction time	5

It should be noted, that the swing weights are dependent on the scales being used for scoring as well as the intrinsic importance of the criteria. This means that it is not possible to assign swing weights until the scales for each criterion have been defined. If an intrinsically important criterion does not differentiate much between the alternatives (the maximum and minimum points on the value scale correspond to similar levels of performance) then that criterion may be ranked quite low.

Due to the special characteristics of criterion C9 (the BCR based criterion) the participants chose to deal with it using a direct rating procedure, as it was important to them to deal with a specific weight-set between the two types of assessments: CBA and MCDA. Thus, discussion led to that C9 was assigned with a relative weight on 0.5 and that the sum of the weights for C1 – C8 should correspond to a relative weight on 0.5 as well. Hence, the CBA and MCDA were assumed to be equally important in the composite assessment.

3.5 Step five – Results and validation

After deriving scores for the alternatives within all criteria and deriving weights for the criteria it was possible to compute a result based on the efforts of the participants at the decision conference; the resulting values are shown in Table 5. It should be noted that the values are calculated using the simple additive value function method.

Table 5. Results of the composite analysis

	C1	C2	C3	C4	C5	C6	C7	C8	C9	Value
Original weights	*90*	*80*	*100*	*60*	*70*	*20*	*60*	*5*		
Normalised weights	*0.09*	*0.08*	*0.10*	*0.06*	*0.07*	*0.02*	*0.06*	*0.01*	*0.50*	
R	1	0	35	0	0	11	0	39	100	**54.15**
BS	100	18	35	100	100	5	100	0	58	**63.10**
BL	70	25	100	51	100	0	100	0	17	**44.05**
G	0	100	0	15	55	100	2	100	0	**15.82**

The results in Table 5 clearly depict that alternative BS is the most attractive followed by alternative R, BL and finally G. If the participants at the decision conference did not agree with this result, or if they were unsure about how some of their assessments had influenced the result, it would be possible to go back to step three or four and revise the assessments or perhaps test different weight-sets for the criteria. However, in the current case study the participants were confident about the result as it corresponded with their own gut feelings, and nobody felt a need to revise the assessments. On the contrary the participants felt that the procedure at the decision conference only confirmed what they already knew, but which they could not present documented arguments for. The procedure using the C-DSS, however, provided them with a means for stating the best solution seen from a combined socio-economic (CBA) and strategic (MCDA) point of view.

4. Discussion

Decisions made in consensus at a decision conference seem to have a fairly higher probability for being implemented than results from a complex decision analysis that only involves one decision-maker who later has to justify his decision for other people (e.g. in a organisation or to the public). Moreover, decisions made by such groups have better terms for working in practice as they have the group's commitment. However, there is one large question that has to be answered: Are decisions made in consensus at a decision conference using a C-DSS more or less valid than assessments and solutions made without these aids? According to Phillips (2006) this is not necessarily the case, however, it is evident that a decision conference provides some advantages regarding: better communication between groups, a common understanding of strategic objectives and hence common commitment towards the objective, improved teamwork, better knowledge and relation to various uncertainties, and finally and foremost decisions that can be defended.

Customising a DSS to fit the specific assessment task in hand places high demands on the decision analysts to indentify and use the most appropriate techniques. In the case study the decision conference was attended by persons with expert knowledge about the decision problem. Consequently, it was possible to make use of a demanding technique like swing weights. In other cases where the participants may have a lower level of knowledge about the decision problem and the criteria to be assessed, a more easily accessible technique – for instance pair wise comparisons or surrogate weights based on rankings (Roberts and Goodwin, 2002) – could be more appropriate. The pair wise comparison technique – which is used on attribute level (scoring of alternatives) in the case study – is not a demanding technique to apply and does not require expert knowledge. However, it is a very appropriate technique when dealing with decision problems in a "local" system, where the task is to identify the best alternative for a given project – not to compare projects of different types or with geographical different locations. If projects are to be assessed on a more "global" level, perhaps comparing projects across borders, the need would arise for a completely different scaling technique such as e.g. direct rating. Hence, different parameters are determining which combination of techniques to apply to the specific decision problem. Setting up the C-DSS is therefore a crucial task for the decision analysts as techniques that are appropriate for one decision problem might be inappropriate for another problem; thereby the participants input will not be treated in a suitable manner and the assessment will in worst case provide misrepresenting results.

5. Conclusions

A combination of techniques for deriving and modelling decision-maker/stakeholder preferences has been introduced applying the process of a decision conference. The decision conference has been proposed to be set up as a five-step procedure leading the participants through the decision process in an easily accessible and transparent manner. The assessment techniques applied are used for developing C-DSS which can support the appraisal of large transport infrastructure projects in need of combining conventional cost-benefit analysis (CBA) with multi-criteria decision analysis (MCDA) to obtain a more comprehensive appraisal featuring all relevant impacts.

It is exemplified through the case study that the proposed C-DSS in combination with the proces of a decision conference is an effective decision aid when complex decisions regarding transport infrastructure projects have to be made. This type of decision problem involves a high number of different stakeholders and decision-makers and a structured process capturing all aspects of the issue is therefore needed. In this respect the proposed methodology provides a customised process which seeks to give everybody an opportunity to express their preferences and influence the outcome. Hence, the C-DSS makes it possible for the decision-makers to make a more informed decision than would be the case if a decision conference were not carried out.

References

Belton, V. and Stewart, T. J. (2002). *Multi Criteria Decision Analysis: An integrated approach.* Kluwer Academic Publishers, London, United Kingdom.

EUNET (2001). *EUNET/SASI, Final Report – Executive Summary.* 4th RTD Framework Programme of the European Commission.

Goodwin, P. and Wright, G. (2004). *Decision Analysis for Management Judgment.* Third Edition, Wiley, Chichester, United Kingdom.

Hiselius, L., Barfod, M.B., Leleur, S., Jeppesen, S.L., Jensen, A.V. and Hjalte, K. (2009). Helhetsorienterad Utvärdering av Kollektivtrafikåtgärder. University of Lund, Sweden.

Jeppesen, S. L. (2009). Sustainable Transport Planning – A Multi-Methodology Approach to Decision Making. Ph.D. thesis, Department of Transport, Technical University of Denmark.

Keeney, R.L. (1992). *Value-Focused Thinking – a Path to Creative Decisionmaking.* Harvard University Press, Massachusetts, United States of America.

Keeney, R.L. and Raiffa, H. (1993). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* Cambridge University Press, New York, United States of America.

Leleur, S. (2000). *Road Infrastructure Planning: A Decision-Oriented Approach.* Second edition, Polyteknisk Press, Kgs. Lyngby, Denmark.

Lootsma, F.A. (1992). The REMBRANDT system for Multi-criteria Decision Analysis via pairwise comparisons or direct rating. Report 92-05 from Faculty of Technical Mathematics and Informatics, Delft University of Technology, Netherlands.

Olson, D. L., Fliedner, G. and Currie, K. (1995). Comparison of the REMBRANDT system with the analytic hierarchy process. *European Journal of Operational Research* 120, pp. 522-539, Elsevier.

Phillips, L.D. (1984). A Theory of Requisite Decision Models. *Acta Psychologica* 56, pp. 29-48, Elsevier.

Phillips, L.D. (2006). Decision Conferencing. A working paper LSEOR 06.85, Operational Research Group, Department of Management, London School of Economics and Political Science, United Kingdom.

Phillips, L.D. and Bana e Costa, C. (2005). Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. A Working Paper LSEOR 05.75, Operational Research Group, Department of Management, London School of Economics and Political Science, United Kingdom.

Roberts, R. and Goodwin, P. (2002). Weight Approximations in Multi-attribute Decision Models. *Journal of Multi-Criteria Decision Analysis* 11, pp. 291-303, Wiley.

von Winterfeldt, D. and Edwards, W. (1986). *Decision Analysis and Behavioural Research.* Cambridge University Press, New York, United States of America.

Paper 4

An MCDA approach for the selection of bike projects based on structuring and appraising activities

Michael Bruhn Barfod
Department of Transport, Technical University of Denmark

Paper published in *European Journal of Operational Research* 218, Issue 3, pp. 810-818 (Elsevier), 2012 – ISI-indexed

Keywords: Decision analysis, problem structuring, multi-criteria decision analysis, REMBRANDT, decision support systems

Abstract

This paper presents an MCDA approach for the structuring and appraising activities of a large and complex decision problem. More specifically, the paper makes use of the three-step structuring process for decision analysis proposed by von Winterfeldt and Edwards: 1) identifying the problem; 2) selecting an appropriate analytic approach; and 3) developing a detailed analytic structure. For illustration of the approach a case study dealing with the assessment task of prioritising and selecting initiatives and projects from a public pool with limited funds is examined throughout the paper. The process is embedded in a decision support system (DSS) making use of the REMBRANDT technique for pair wise comparisons to determine project rankings. A procedure for limiting the number of pair wise comparisons to be made in the process is in this connection presented. Finally, strengths and weaknesses in the approach are discussed and conclusions are made.

1. Introduction

When making decisions, decision-makers (DMs) will in most cases try to choose the optimal solution. Unfortunately, a true optimal solution only exists if you are considering a single criterion. In most real decision situations, basing a decision solely on one criterion is, however, insufficient. Probably several conflicting and often non-commensurable objectives should be considered. As a result of this it is impossible to find a genuine optimal solution, a solution which is optimal for all DMs under each of the criteria considered (Løken, 2007). Multi-criteria decision making (MCDM) is a generic term for methods that assist people in making decisions using their own preferences in cases where more than one conflicting criterion exists. Using MCDM can be said to be a way of dealing with complex problems by breaking them into smaller pieces. After weighting procedures and judgments of the smaller components the pieces can be reassembled to present an overall picture to the DMs.

Another term used instead of MCDM is multi-criteria decision analysis (MCDA), where the use of 'analysis' instead of 'making' emphasises that the method should assist the DMs in making decisions (as the method itself cannot make the decision). Hence, the aim of MCDA is to assist the DMs to choose, rank or sort alternatives within a finite set according to two or more criteria so that they feel comfortable with the final decision (Chen et al., 2008). By using MCDA the DMs should feel that all important criteria have been properly accounted for, which should help to reduce the possibility of post-decision regret (Belton and Stewart, 2002). Ideally, the MCDA method will help the DMs to understand and identify the fundamental criteria in the decision problem and avoid making important decisions only out of habit.

Structuring the decision problem – taking it from an initially vague and ill-defined problem to one that can be formulated, modelled and analysed mathematically – is by von Winterfeldt and Fasolo (2009) stated to be the hardest yet most crucial part of an operations research (OR) analysis. This is a focus of decision analysis, where the emphasis of problem structuring is on shaping general statements by the DMs about their goals, concerns, issues and uncertainties and turning these statements into a clear and transparent representation of the decision problem which can be mathematically formalised using the principles of decision theory, see e.g. von Winterfeldt and Edwards (1986, 2007) and Belton and Stewart (2002).

This paper presents the structuring and appraising activities for the public Danish pool for more bike traffic, which was conducted in late 2009 as consultancy for the Danish Road Directorate. The bike pool is a result of a political agreement concerning a new green profile for traffic planning in Denmark supporting bike projects with 1 billion DKK in the period from 2009 to 2014. As a part of the political agreement 150 million DKK was in 2009 allocated to support initiatives and projects (onwards referred to as projects) that contributes to make bikes a more attractive means of transportation. The aim of the pool was to move users from car traffic, but also public transportation, to bikes. The bike pool was open for applications of widely varying characters, and in principle it was possible for everybody to apply for subsidies from the pool. As a result of this a total of 133 project applications were submitted from municipalities, regions, organisations, companies and research institutions. The projects amounted to a total sum of approximately 1 billion DKK, which corresponded to a subsidy sum of approximately 450 million DKK (most projects were eligible for between 30 and 50 % subsidy and a few projects for 100 % subsidy from the pool). Hence, there was a need for an appraisal of which projects should be given subsidies from the pool, as it was impossible to give subsidies to all the projects. The technical evaluation task was henceforth to design and apply a series of principles and methods which were capable of handling this large quantity of projects in an appropriate and optimal way. This, so that the total means of the pool could be allocated to those projects and initiatives that contributed the most to the overall objective.

In Denmark it is a basic point of view that appraisals of transport projects shall be based on socio-economic evaluation to state if the projects are economically feasible or not. This is normally conducted using a manual for socio-economic appraisal from 2003 (Danish Ministry of Transport, 2003) and the newest edition of traffic economic unit prices (the key figures' catalogue). However, currently no such foundation exists for economic appraisals of bike projects, and moreover it was impossible to conduct impact calculations on the applications submitted due to their vaguely written form and content. As the assessment task went beyond socio-economic calculations and as the limitations of the task (time constraints, budget limitations etc.) made it impossible to set out a foundation for this, it was decided to use a methodological approach which was based on principles for value measurement different from traditional cost-benefit analysis (CBA). Hence, the concept of MCDA was introduced to deal with the assessment task in order to ensure an appropriate and comprehensive assessment, while at the same time making it possible to perform the appraisal within a limited time frame. Thus a decision support system (DSS) named the CPP-DSS (CykelPuljePrioritering (Danish for Bike Pool Priority)) was developed. The DSS was based on a qualitative evaluation, but with a perspective saying that the approach to be applied could be based on a combined use of CBA and MCDA as it is e.g. described by Leleur et al. (2007) and Barfod et al. (2011).

With reference to the previous work on decision analysis conducted by other researchers this paper deals with three main research questions: Can the theory of decision analysis be useful to structure a decision problem involving a large number of options, multiple objectives and multiple stakeholders? Can the appraisal of a decision problem using MCDA be operationalised into a DSS that can inform the DMs in terms of both interaction and interpretation of the results? And finally, can a set of appropriate guidelines be formulated for the appraisal of widely varying projects using the DSS?

This paper is organised as follows. After this introduction a literature review on structuring decision problems for OR in general and decision analysis in particular is conducted. In the following three sections a process for structuring and appraising a decision problem is conducted on the case study comprising the three steps of: identifying the problem, selecting an analytical approach and developing a detailed analytical structure. Finally, conclusions are made and perspectives for the future modelling work are given.

2. Problem structuring using decision analysis

At the most basic level a decision analysis structure defines the scope of a decision problem, including the DMs and stakeholders, their values and alternatives, the range of consequences of concern, and the key uncertainties (von Winterfeldt and Fasolo, 2009). Scanning the literature on structuring problems for decision analysis it is found that structuring does not only involve framing the problem, but also two additional steps of selecting an appropriate structure and developing this in details before numerical modelling and analysis begins (von Winterfeldt and Edwards, 1986, 2007; Keeney, 1992; Belton and Stewart, 2002; Goodwin and Wright, 2009). In this respect problem structuring methods (PSM) can be very helpful to support groups in confronting the three steps (Mingers and Rosenhead, 2004).

There is much to be learnt about problem structuring from the body of work stemming from the fields of what is collectively referred to as "soft" OR or PSM, see Rosenhead and Mingers (2001). Under this are among others the following approaches, which pay attention to multiple objectives and multiple perspectives in a more or less formal way: Strategic Options Development and Analysis (SODA) by Eden (Eden and Ackermann, 2001), and more recently extended to the concept of journey making (Ackermann and Eden, 2001); the Strategic Choice Approach by Friend and Hickling (Friend, 2001) and the Soft Systems Methodology (SSM) by Checkland (Checkland, 2001). Each of these methods has something to offer problem structuring for MCDA, see e.g. Neves et al. (2009) using SSM for structuring a MCDA model.

Phillips (1984, 2007) deals with the concept of a "requisite decision analysis model" which he defines as one that is sufficient in form and content to resolve the issue at hand. Moreover, he states that a decision model is requisite if no new intuitions arise in the group. While requisite modelling can be best recognised when a full model is developed, including elicitation of data, this notion can also be applied to decision analysis structure, implying that there can be structural representations that are simple enough to capture the essence of a decision problem, and no more complicated than necessary to obtain sound insights. A decision analysis structure is thus requisite if no additional insights emerge that will lead to significant additions or modifications of the structure (von Winterfeldt and Fasolo, 2009).

MCDA is deemed to offer a sound methodology for promoting a good decision making process (Keeney and Raiffa, 1993) and the field is characterised by a variety of different techniques and approaches (Stewart and Losa, 2003). A representative excerpt of the literature on decision analysis (von Winterfeldt and Edwards, 1986, 2007; Keeney, 1992; Keeney and Raiffa, 1993; Belton and Stewart, 2002; Goodwin and Wright, 2009) indicates the relevance of distinguishing between the following eight different analytic structures depending on the type of the problem being either a multi-attribute evaluation problem, or a decision problem involving significant uncertainties, or a probabilistic inference problem:

- Evaluation problems
    - Means-ends networks
    - Objectives hierarchies
    - Consequence tables
- Decision problems under uncertainty
    - Decision trees
    - Influence diagrams
- Probabilistic inference problems
    - Event trees
    - Fault trees
    - Belief networks

First, almost all problems have multiple objectives and thus some structuring of alternatives and objectives is always useful (Keeney, 1992). Simple objectives hierarchies and consequence tables help to clarify the key relationships between alternatives and objectives. If data concerning consequences are not readily available, ranking projects by objectives can be illuminating. Second, decision trees are useful, if there are clear, important, and discrete events that stand between the implementation of the alternatives and the eventual consequences. Decisions, for example, dealing with major disasters, terrorism, and the like lend themselves to decision trees. The multiple consequence part of this type of problem can be handled by listing all consequences at the end of the decision tree and determining an equivalent value or utility through standard multi-attribute utility analysis (Keeney and Raiffa, 1993). Influence diagrams are most useful when some of the uncertain variables are continuous and causally linked. In this case it may be easier to develop a deterministic model that calculates the propagation of causal effects and then to superimpose a probabilistic simulation to assess the overall uncertainties (von Winterfeldt and Edwards, 2007). Multiple objectives aspects can easily be integrated with influence diagrams. Fault trees, event trees, and belief nets are special to inference problems.

Decision trees and diagrams are the major structuring tools of decision analysis. However, building such trees and diagrams can be regarded as a fairly specific activity. In order to deal with the structuring within a broader perspective von Winterfeldt and Edwards (1986, 2007) proposed the following three step procedure for structuring decision problems in a decision analysis: 1) identify the problem, 2) select an analytical approach, and 3) develop a detailed analysis structure. Although these steps seem reasonably distinct, the intellectual work behind them can be extremely recursive. The decision analyst should prepare himself to go through each step several times and probably restructure the problem a number of times.

It can be noted that PSMs structuring allows local, partial solutions rather than global solutions that imply a merging of different views (Mingers and Rosenhead, 2004). This means that values and uncertainties are structured in qualitative incommensurable form and it distinguishes the structuring of PSMs from the structuring activities in decision analysis, which aims at developing a quantitative model of the DMs values (multi-attribute utility problems) and perceptions of uncertainties (uncertainty problems).

In the following sections the three steps of von Winterfeldt and Edwards (1986, 2007) will be applied to the case study of distributing subsidies from the Danish bike pool. The focus is on the process and the techniques applied; the results are the property of the Danish Road Directorate and will thus not be specified. However, this should not affect the scope of the paper.

3. Identifying the problem

When they are first encountered, some decision problems appear to be overwhelmingly complex. Any attempt at clear thinking can be frustrated by the large number of interrelated elements that are associated with the problem, so that, at best, the unaided DMs can have only an unclear perception of the issues involved (Goodwin and Wright, 2009). Thus, when DMs approach a decision analyst for consultancy with a problem, they will often only have a general idea of what the problem is, and in many cases the initial discussions between the DMs and the decision analyst change the character of the problem.

In identifying the extent of the problem von Winterfeldt and Edwards (1986, 2007) suggest that it is useful to answer five simple questions, see Table 1. For this purpose a preliminary meeting was held with key persons from the Danish Road Directorate in order to answer the questions. In addition to this a literature study was conducted in order to expose previous research within the field of appraisals of bike projects. The questions and their answers are outlined in Table 1.

Table 1. Questions identifying the problem

Question	Answer
What is the nature of the problem?	At the initial stage the bike problem appeared to be relatively straightforward. However, studying the background for doing appraisals for bike projects made it clear that an alternative appraisal methodology than conventional CBA is needed as only very limited research has been made on socio-economic impacts of bike projects.
What is the problem environment and who are the stakeholders?	Many different stakeholders have an interest in the decision problem. Among these the Danish Bicycle Union (DCF), the United Danish Motorists (FDM) and the Municipalities in Denmark can be mentioned as the most important ones.
Who are the DMs and what are their values?	The final decision making lies with the political parties which made the agreement regarding the origin of the bike pool. However, the appraisal is conducted with the Danish Road Directorate as the DMs, who then make an impartial technical recommendation for the political parties to consider.
What is the purpose of the analysis?	To support the Danish Road Directorate in choosing the projects that contributes the most to the pool's overall goal of making bikes a more attractive means of transportation.
What is the generic class of alternatives?	The projects differ in their descriptions and types, and it is for this reason found to be necessary to divide them into different groups with specific characteristics. This contributes to make the initial part of the analysis simpler.

Answering the questions in Table 1 made it clear that the decision problem was of a complex character involving many different stakeholders whose preferences needed to be accounted for in the analysis. Even though the final decision regarding the case study was in the hands of the political parties, government officials from the Road Directorate were to act as DMs for practical reasons.

As previously mentioned, the bike pool is a result of a political agreement concerning a new green profile for traffic planning in Denmark where the aim is to move users from car traffic, but also from public transportation to bikes. The pool was open for projects of a widely varying character, and in principle it was possible for everybody to apply for subsidies from the pool; as a result of this a total of 133 project applications were submitted from municipalities, regions, organizations, companies and research institutions. The prerequisites of the applications, which were almost non-existing, resulted in many different types of projects with highly varying impact descriptions. Some applications were well-described and well-defined with regard to expected impacts for the projects, while other applications more or less only consisted of a map with an indication of where to build a bike path. For this reason a division of the projects into different types needed to be made.

4. Selecting an analytical approach

At some stage the emphasis of the analysis needs to move from problem structuring to model building where a specific analytical approach is used for the development of a framework for the appraisal of the projects. Model building should in this context be regarded as a dynamic process, informed by and informing the problem structuring process, and interacting with the process of appraisal. It may involve some iteration, search for new criteria, discarding, reinstating and redefining old ones, and further extensive discussions amongst the participants in the process. Moving from a broad description of the problem, whether it is a simple clustering of ideas, a fully elaborated map, or some other representation of the issue, to a preliminary definition of a model for MCDA, requires a good understanding of the chosen approach to multi-criteria modelling. The nature of the model which is sought will differ according to the nature of the assessment task, whether alternatives are explicitly or implicitly defined, and the particular approach selected for the analysis.

For the bike problem, which can be characterised as a multi-attribute evaluation problem, an initial workshop with the participation of key stakeholders (see Table 1) and DMs was held with the purpose of identifying a set of fundamental objectives, see Keeney (1992). The workshop was set out as a 'futures workshop'; see e.g. Leleur (2008), where an impartial facilitator led the participants through the three phases of 'criticising', 'fantasising' and 'implementing'. By doing this a long list of more or less relevant objectives was identified. However, many of the objectives identified were important as they served other ends objectives. By pursuing the means-ends chain of objectives one will eventually arrive at objectives, whose importance are self evident – these are the fundamental objectives (Keeney, 1992). Interviews or workshops to elicit objectives typically generate many objectives that are not fundamental to the decision problem (Belton and Stewart, 2002). These falls into two categories: means objectives and process objectives (von Winterfeldt and Edwards, 2007). Means objectives can affect fundamental objectives (e.g. air pollution is a means to create health impacts), but are not themselves relevant for the assessment. Process objectives on the other hand are those that refer to the process by which a decision is made rather than its consequences. Stakeholders and DMs expect all objectives to be represented in some form in an assessment. Process objectives, however, should be separated out and considered when designing and implementing the analysis process (Keeney, 1992). During the workshop focus was on eliciting operational objectives. However, checks were subsequently performed on selected test-projects in order to assure that also the criteria of completeness, absence of redundancy and decomposability were met. This led to minor adjustments of the objectives. Table 2 shows the objectives that were elicited at the workshop for the appraisal of bike projects on an overall level. The fundamental objectives are shown in the left column and the selected measures in the right column.

Table 2. Objectives and measures for evaluating bike projects on an overall level

Objectives (fundamental)	Measures (selected)
Mode choice	An estimate of the number of new-coming bikes on the roads as a result of the project.
Accessibility	An estimate of how much the congestion will be reduced as a result of the project (measured in hours).
Safety	An estimate of the reduction in injuries and accidents as a result of the project.
Perceived risk	An estimate of how the perceived risk will be reduced on given stretches as a result of the project (interviews with users should be conducted to cover this).
Practicability	An estimate of how visible and useable the project will be for the users (interviews with users should be conducted to cover this).
Costs	Measured in Danish Kroner (DKK)

Structuring the objectives does not lead directly to a formal analysis. Instead, it clarifies for the DMs and stakeholders how their concerns are being handled. A MCDA method can then be used to summarise the consequences of projects and to aggregate these to a common value metric. Thus, when choosing a MCDA approach, there are many aspects to consider. The most important is to apply an approach that really measures what it is supposed to measure (validity). Different approaches are likely to give different results, so an approach that reflects the DMs' 'true values' in the best possible way should be chosen. In addition the approach should provide the DMs with all the information they need, and it must be compatible with the accessible data (appropriateness). Moreover, the approach must be easy to use and understand (Hobbs and Meier, 2000). If the DMs do not understand what is happening inside the methodology, they will perceive it as a 'black box'. The result of this may be that the DMs do not trust the recommendations from the methodology. In that case it is meaningless to spend time applying the approach. Some guidelines for choosing an appropriate MCDA method can be found in e.g. Guitouni and Martel (1998) or Zopounidis and Doumpos (2002).

The DSS developed for the CPP case was subsequently specified to be based on a value measurement model. Using this approach a numerical score is assigned to each project producing a preference order such that project *j* is preferred to project *k* if and only if *V(j) > V(k)*. The various criteria are given weights that represent their partial contribution to the overall score, based on how important the criteria are to the DMs. Ideally, the weights should indicate how much the DMs are willing to accept in the trade-off between two criteria (Keeney and Raiffa, 1993; Belton and Stewart, 2002).

During the previously mentioned workshop the participants expressed the expectation that the assessment technique to be used in the further analysis should be easy to understand both for professionals and non-professionals. Moreover, the judgments to be made should be as simple as possible in order to ensure transparency both in the process, but also when the results afterwards need to be defended. Due to this, the character of the objectives to be measured, and the information level of the projects the CPP-DSS was chosen to be based on a technique using pair wise comparisons for the elicitation of scores and weights. The usefulness of this technique is, in respect of the case study, that the decision problem is decomposed into a hierarchy of more easily comprehended sub-problems, each of which can be analysed independently. The elements of the hierarchy can relate to any aspect of the decision problem – tangible or intangible, carefully measured or roughly estimated, well- or poorly-understood – anything at all that applies to the decision at hand. Once the decision hierarchy is built, the DMs can systematically evaluate its various elements by comparing them to one another two at a time, with respect to their impact on an element above them in the hierarchy. In making the comparisons, the DMs can use concrete data about the elements, or they can use their judgments about the elements' relative meaning and importance. It is the essence of the technique that human judgments, and not just the underlying information, can be used in performing the evaluations.

Thus, the CPP-DSS makes use of the REMBRANDT (Ratio Estimation in Magnitudes or deci-Bells to Rate Alternatives which are Non-Dominated) technique by Lootsma (1988, 1999) which is a multiplicative version of the Analytic Hierarchy Process (AHP) by Saaty (1977, 2001). REMBRANDT proposes to overcome three issues regarding the theory behind AHP, for elaborations of these see Olson et al. (1995), Ramanathan (1997), Lootsma (1999), and van den Honert and Lootsma (2000). The technique makes use of pair wise comparisons between projects to determine subjective impacts under each criterion in the assessment and between criteria in order to determine their relative importance. The DMs pair wise comparative judgment of the projects P_j versus P_k is captured on a category scale to frame the range of possible verbal responses. This is converted into an integer-valued gradation index δ_{jk} according to the REMBRANDT scale in Table 3.

Table 3. The REMBRANDT scale (van den Honert and Lootsma, 2000)

Comparative judgment	Gradation index δ_{jk}
Very strong preference for P_k over P_j	-8
Strong preference for P_k over P_j	-6
Definite preference for P_k over P_j	-4
Weak preference for P_k over P_j	-2
Indifference	0
Weak preference for P_j over P_k	+2
Definite preference for P_j over P_k	+4
Strong preference for P_j over P_k	+6
Very strong preference for P_j over P_k	+8

Thus there are five major, linguistically distinct categories in Table 3: indifference, weak, definite, strong and very strong. Moreover, there are four so-called threshold categories between them which can be used if the DMs are in-between the neighbouring qualifications. The values obtained by the comparisons are gathered in a comparison matrix and using the principles of the REMBRANDT technique relative project scores are calculated (this procedure is also applied on criteria level). Finally, scores are aggregated by the product of projects relative scores weighted by the power of weights obtained from the analysis of hierarchical elements above the projects. This leads to a final score for each project, which allows a subjective rank ordering of the projects.

It should be noted that the problem could benefit from other MCDA approaches as well. The original AHP could for instance be applied in the same way as the REMBRANDT technique, but outranking procedures, see e.g. Roy and Vanderpooten (1996), could also be an option. The REMBRANDT technique was chosen due to the simplicity and transparency of the pair wise comparisons which are easy to explain and make use of; this was a very important issue for the DMs and stakeholders. Compared with the original AHP, REMBRANDT was chosen due to the theoretical improvements (Olson et al., 1995).

5. Developing a detailed analytical structure

At this point step 1 and 2 of von Winterfeldt and Edwards (1986, 2007) process have been dealt with. Therefore this section will concentrate on developing a detailed structure for the assessment of the case problem. This concerns the task of distributing subsidies from a public pool to bike project applications in order to make bikes a more attractive means of transportation for which the pair wise comparison technique REMBRANDT has been chosen.

In consultancy with the DMs the 133 projects were first divided into 3 main pools according to a characterisation of the project type. It was found convenient to express these as: innovation, safety or bike city projects. Second, a screening of the projects' characteristics shown it necessary to divide the 3 main pools into 9 sub-pools (3 for each main pool) describing each project type in a higher level of detail. As depicted in Figure 1 the sub-pools contained: knowledge/research projects, plan and concept projects and campaign projects (the innovation pool); school road projects, bike path projects and bike tourism projects (the safety pool); and bike city projects, commuting projects and bike parking projects (the bike city pool).

Consequently, the prioritisation task consisted of producing prioritised lists for each of the 9 sub-pools and hereafter combined pool lists for each of the 3 main pools. The final prioritised lists (one for each of the 3 main pools) thus contained the projects which were found worthy for subsidies from the bike pool. It should be noted, that it was by request from the DMs to divide the projects into 3 pre-specified main pools, and to derive a prioritised list for each of these instead of one overall list.

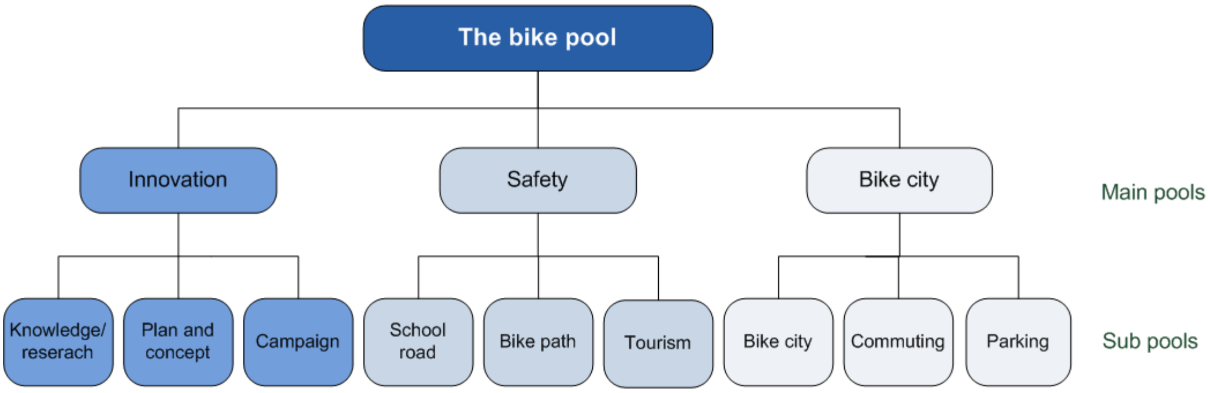


Figure 1. Division of the projects into different types

It was moreover necessary to determine what impacts (here denominated as criteria) that characterised the different project types in order to assess how well the projects contributed to promote the overall goal of the bike pool. In the previously mentioned workshop, with participation of key stakeholders and DMs, a long list of possible criteria had been created on the way to define the objectives in Table 2. In consultancy with the DMs relevant criteria from this list were now assigned to the sub-pools ensuring operationality. Each of the 9 sub-pools was henceforth based on a set of criteria defined to measure a change in impact relevant for the specific type of project. It was decided to limit the number of criteria to a maximum of the four most important for each sub-pool although more criteria were identified. This was done as a consequence of time constraints for the assessment process, and was moreover based on the assumption that more criteria would not contribute significantly to the assessment (von Winterfeldt and Fasolo, 2009). Table 4 provides an overview of the criteria used for comparisons in the sub-pools. For comparisons of projects across the sub-pools the objectives from Table 2 are used as general criteria.

Table 4. List of criteria

The innovation pool			The safety pool			The bike city pool		
Know-ledge/ research	**Plan and concept**	**Campaign**	**School road**	**Bike path**	**Tourism**	**Bike city**	**Commuting**	**Parking**
-Innovation	-Facilities	-Relevance	-Perceived risk	-Coherence in the network	-Coherence in the network	-Mode choice	-Time savings	-Coheren-ce with terminals
-Possibilities	-Mode choice	-Visibility	-Coherence in the network	-Perceived risk	-Experien-ces along the route	-Coherence in the network	-Coherence in the network	-Service level
-Communi-cation	-Innovation	-Possibili-ties	-Behaviou-ral change	-Priority in relation to other modes	-Service level on the paths	-Coherence with terminals	-Mode choice	-Aesthe-tics
		-Behaviou-ral change		-Time savings	-Perceived risk	-Perceived risk	-Perceived risk	

As a result of the characteristics of the case problem and the preferences of the DMs the overall assessment task was structured as depicted in Figure 2 where the overall goal is based on the 9 different project types in the sub-pools. Each of the 9 sub-pools is based on a set of criteria, where each criterion measures a change in impact relevant for the specific sub-pool, see Table 4. The criteria sets are depicted on the following level in Figure 2, and the bottom level shows the projects. The projects are divided into sub-sets for each sub-pool placing projects with similar investment costs in the same set in order to obviate the cost criterion at the initial stage. This exercise is done as a consequence of the large sizes of some of the sub-pools which made it practically inconvenient to use pair wise comparisons. By dividing the sub-pools into sub-sets considerably fewer pair wise comparisons have to be made at the initial stage. At this stage the costs are neutralised as a decision factor but the cost criterion becomes relevant again when projects from different sub-sets are compared. Pair wise comparisons of the projects are conducted for each sub-set under each of the criteria. The results of these pair wise comparisons are an assessment of the relative performance of each sub-set's projects in relation to the criteria applied. It should be noted that Figure 2 only shows the principal structure of the decision problem, in practice the 9 project sub-pools were of a very varying size containing between 4 and 36 projects.

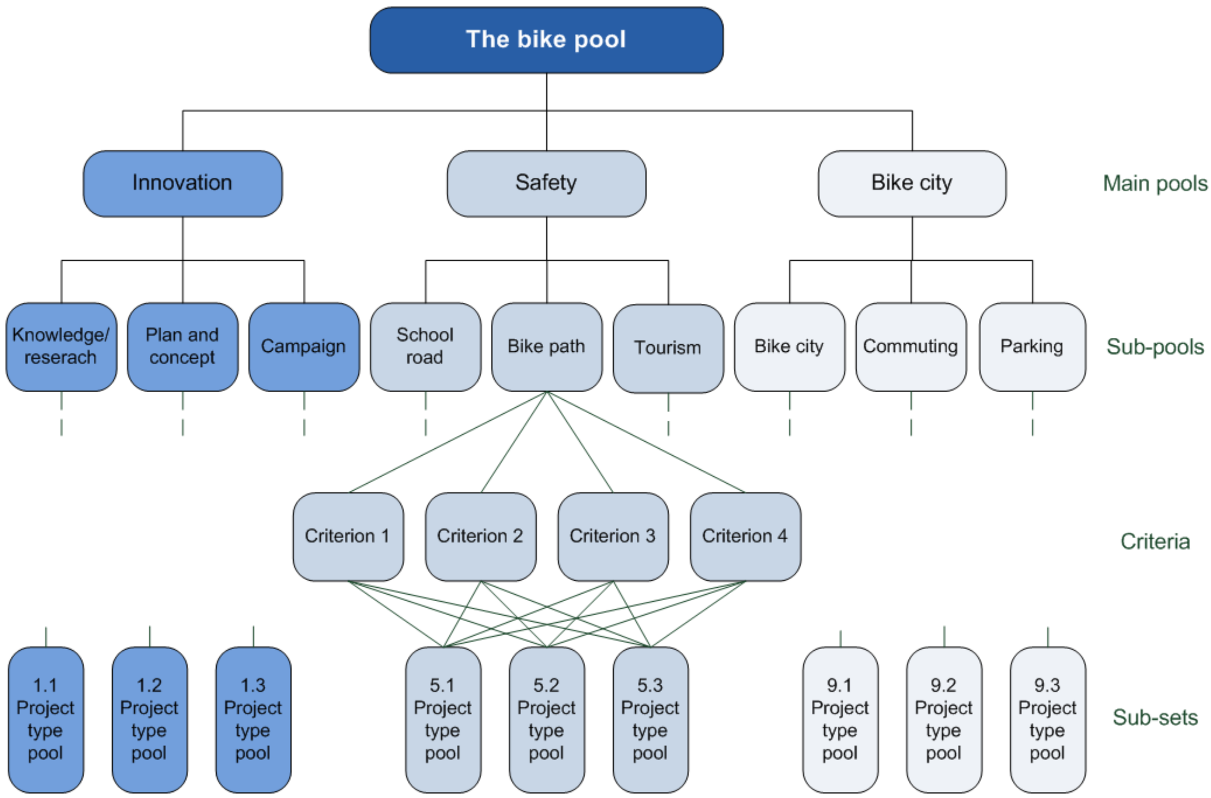


Figure 2. Principal structure of the decision problem

In order to make the assessment of the projects as comprehensive as possible, the DMs met to systematically discuss and analyse the issues at a decision conference as described by Phillips (2007). The objective of such a decision conference is to deal constructively with the conflicting issues at hand so a common understanding of the issue can be achieved, see Mustajoki et al. (2007). In fact several decision conferences were held as the size of the decision problem made it impossible to handle everything in a two to three day meeting. Instead single day conferences focussing on specific sub-pools were held. This made it possible for the DMs to reflect over previous assessments in the days between the conferences and revise some judgments if this was felt appropriate. The conferences were held at the Technical University of Denmark’s property in order for the DMs to get out of their normal surroundings as recommended by Phillips (2007) thereby making it easier for them to concentrate on the relevant issues. The conferences were controlled by an impartial facilitator that guided the DMs through the process of deriving input to the decision model – the CPP-DSS – which was operated by a decision analyst. Thus the CPP-DSS functioned to model the viewpoints of the participants and to appraise the projects in a manner that could be accounted for afterwards.

In this context it should be noted that all choices and assessments made during these decision conferences were documented in a so-called assessment protocol, which had the purpose of creating a documented rationale for later justifications. Efforts were made for the participants to reach consensus on each of the comparisons before moving on to the next. In the cases where it was not possible to agree upon the comparisons the different viewpoints were noted with a view to a later sensitivity analysis if felt needed by the participants.

The participants at the decision conference were by request from the political support base solely selected by the Danish Road Directorate. Thus the group consisted of government officers who all had specific areas of expertise in bike related issues and experience in the use of decision support tools. As a result of this it was possible to keep the analysis on a high level of problem understanding.

In overview the following steps A – F comprise the process of the CPP-DSS:

A. The applications for the projects are scrutinised and those applications that are not satisfying the standards defined are sorted out. The remaining projects are organised in the 9 sub-pools describing specific project types.

B. The projects are in each of the 9 sub-pools further divided into sub-sets consisting of 3 to 6 projects with similar project costs (in all 24 sub-sets are made).

C. By use of the REMBRANDT technique each sub-set is assessed using pair wise comparisons. Pair wise comparisons are made of all projects under all criteria in the sub-sets' criteria set. Next, pair wise comparisons are made of the criteria in order to determine the criteria weights. Following, project scores are derived and used to rank the projects in the sub-set. The result is 24 prioritised lists in total.

D. The highest ranked projects in each sub-set are gathered in a new set under the sub-pool. These so-called number ones's are compared pair wise under the same criteria as in step C, however, the cost criterion is now added as the projects have different costs. Consequently, project scores are again calculated and the projects are ranked within their sub-pool. The result is 9 prioritised lists.

E. It is now possible to create 3 prioritised lists using the so-called general criteria depicted in Table 2. From each pool-list the projects recommended for subsidy are identified using pair wise comparisons under the general criteria. These general criteria ensure that comparisons are made with regard to the overall goal of the bike pool. The projects are picked out one at a time as one of them reaches the highest rank on the list. When the budget frame is empty the process stops.

F. The 3 prioritised lists are presented for the political parties behind the agreement regarding the origin of the bike pool to consider.

When a project is picked out for subsidy in the previously mentioned step E, the project which was ranked second best in the original sub-set takes its place. Hence, pair wise comparisons have to be made again in order to determine whether this specific project is better than other projects, which were ranked higher in other sub-sets. This procedure ensures that all projects 'compete' at the same terms across the sub-sets and sub-pools, while at the same time it keeps the number of pair wise comparisons to a minimum. The procedure, however, does not necessarily ensure that all 9 sub-pools or projects with varying costs will be present in the final list. This depends on the specific assessments made during the process.

6. Discussion and findings

The results of the CPP-DSS were three lists – one for each main pool – consisting of those projects found appropriate for a subsidy from the bike pool. Figure 3 depicts the structure of the three lists determined by the analysis. Note that the letters denominates different project ID's, which are not shown here because of duty of confidentiality towards the Danish Road Directorate.

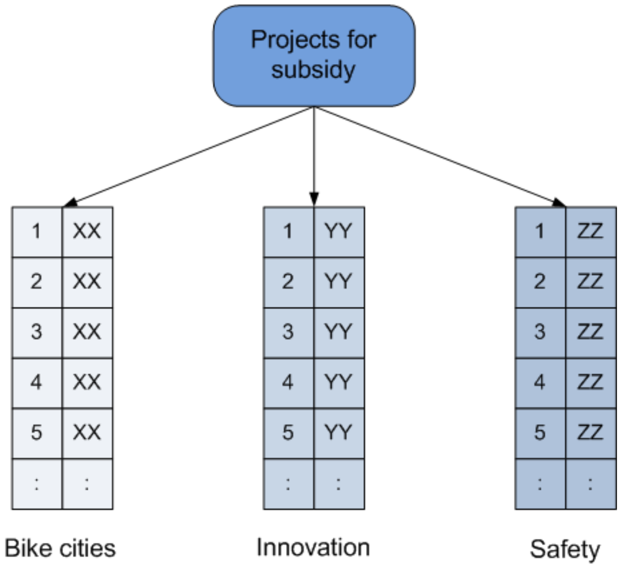


Figure 3. Three prioritised lists of projects are the results of the CPP-DSS

It can be noted that if all 133 projects were to be compared with each other in one large pool under all the criteria used one would have to conduct approximately 300,000 pair wise comparisons; a number which is evidently impossible to handle in practise. If, on the other hand, the projects were divided in 9 project type pools with 15 projects in each and assessed under 5 criteria (the four shown Table 4 plus the cost criterion) one would have to conduct approximately 5,000 pair wise comparisons; which is still a number practically impossible to handle. When using the CPP-DSS with the described analytic structure approximately 600 pair wise comparisons were conducted during the process. This is a large number when each comparison has to be well-argued, but it is manageable over a time period of a few weeks with frequent assessment meetings.

When conducting this type of appraisal the identification and definition of the criteria set is a crucial factor. The criteria need to be defined unambiguously and in a way so that no overlapping takes place. Moreover the group doing the assessments need to have a shared understanding of the criteria to minimise matters of dispute that can arise both during and after the assessment process.

An important issue that needs to be addressed when dealing with these many different types of projects (as has been the case for the CPP-analysis) is the placement of projects in sub-pools. The question addressed is whether a project which performs poorly within its sub-pool might perform better on criteria from another sub-pool. For this reason it is of utmost importance that the projects are placed in the proper sub-pool from the beginning of the analysis. If not, the projects are in risk of not being treated fairly and the funds from the bike pool not distributed in an optimal way. This of course sets some high requirements for the initial work on placing the projects in the right sub-pools.

Another important issue is the assignment of criteria weights as these are determined by individuals using the MCDA method. On the other hand, the performance (scores) of projects for each criterion is determined somehow more objectively, even if artificial scales are used for non-quantifiable criteria. However, the REMBRANDT technique contributes to overcoming this disadvantage by deriving weights in a quasi-independent manner, using pair wise comparisons that make it difficult to promote open biases towards specific criteria. Thus, REMBRANDT is a common method used for prioritisation when having a wide variety of choices. More specifically, with regard to the application of the DSS for the case study, the group that conducted the comparisons was composed of DMs involved in the project.

In addition to the above mentioned some specific findings can be related to the structure of the decision problem:

1. The structuring task should be conducted in close dialogue between the analysts, the DMs and the stakeholders. The dialogue should preferably be highly interactive and iterative leaving options for restructuring during the process.

2. Focus should be on solving the problem, not forcing a particular analytic structure onto the problem. This could even lead to the conclusion that the analytic structure originally considered for solving the problem (for the CPP case: CBA) is not the most appropriate for the problem.

3. A good structure emerges when social and technical facilitation skills are combined. The social skills enhance the likelihood that the DMs and stakeholders will participate in the process, provide important input, and appreciate the results. The technical skills assure that the analysis is logical, simple, manageable, and still relevant to the concerns of the DMs and stakeholders.

4. The CPP-DSS featuring the REMBRANDT technique has shown to be a useful tool when dealing with large complex problems that are in need of a clear structure in order to be solved. The method is moreover easily accessible for the DMs due to the simplicity of the pair wise comparisons.

5. It is seen as a major feature of the CPP-DSS that the various inputs needed from the DMs can help generate important discussions in the group. A future research task will be to explore the modelling and DM interaction further with the purpose of

improving the learning and understanding among the DMs about the actual non-standard appraisal task.

6. Overall, inclusion of socio-economic elements in the assessment of transport projects is widely preferred. Future developments of the CPP-DSS should for this reason work towards inclusion of CBA as the necessary socio-economic foundation become available from research on this topic.

After the finalisation of the assessment task an evaluation meeting was held with the DMs where feedback on the entire process was given. Overall, the DMs were most satisfied with the process agreeing that a requisite model was achieved. The model was perceived to be very useful as it was simple enough to capture the essence of the decision problem, and not too complicated to understand for neither professionals nor non-professionals. The DMs expressed their satisfaction with the approach taken based on several small-scale decision conferences each focussing on specific project types. This enabled the participants to reflect on previous assessments and make corrections at the following decision conference if they felt a need for this. The use of pair wise comparisons was considered appropriate as the problem was decomposed into simple sets of judgments for the participants to consider. The documented rationale in form of the assessment protocol was in this respect a helpful tool both to remind the participants about previous assessments made, but also in informing the politicians and stakeholders about the choices made in the process. The DMs moreover stated that the analysis structure was requisite as no additional insights emerged along the process that led to significant additions or modifications of the structure.

7. Conclusions

This paper has described the structuring and appraising activities associated with a major decision analysis of projects to promote biking activities in Denmark – the CPP problem – and shows that decision analysis using MCDA can be a useful approach for structuring and appraising large and complex decision problems. Specifically, the paper examines the three-step structuring process for decision analysis proposed by von Winterfeldt and Edwards consisting of: framing the problem, selecting an appropriate structure, and developing this in detail before beginning the numerical modelling and analysis. The process has been applied to the CPP decision problem and emphasises the importance of creating a clear analytic structure before attempting to solve the decision problem. The use of decision analysis for the structuring of this specific case problem with a large number of options and multiple objectives was found to be very useful by the DMs. They felt that they had gained sufficient insight in the issues along the process and were for this reason well equipped for defending the results when afterwards facing both politicians and the public.

The MCDA approach presented in the paper has been based on the REMBRANDT technique for pair wise comparisons, which is found relevant as an assessment tool for the specific CPP case study where data and resources were limited for the appraisal. Due to the structure of the decision problem a procedure for limiting the number of pair wise comparisons needed to be made in the process has been set out and applied with a good result. In fact this procedure limits the number of necessary pair wise comparisons from approximately 300,000 in a complete analysis to approximately 600 when using the CPP-DSS. The paper has demonstrated that decision problems such as the treated CPP case do not present themselves in a structured form, complete with lists of alternative courses of action and decision making objectives (criteria), and ready for systematic analysis. Therefore problem structuring in terms of alternatives (projects) and criteria has been a main concern when illuminating the CPP methodology thereby being in accordance with the view that the treatment of components of a problem structure are central to the methodologies of MCDA.

The CPP case is indicative of a number of application areas where approaches similar to the CPP-DSS can be made use of but probably no single best problem structuring method for all decisions exists. The influence of contextual factors on the decision process is significant as is the initial frame taken by the DMs. These and other influences should be studied further as they impact the problem structuring process. The way the CPP-DSS reduces the number of pair wise comparisons can be useful due to its easy transferability to other problem applications generically like the treated.

References

Ackermann, F., & Eden, C. (2001). SODA – journey making and mapping in practice. In J. Rosenhead, & J. Mingers (Eds.), *Rational analysis for a problematic world revisited* (pp. 43-61). (2nd ed.). John Wiley & sons.

Barfod, M.B, Salling, K.B., & Leleur, S. (2011). Composite decision support by combining cost-benefit and multi-criteria decision analysis. *Decision Support Systems*, 51, 167-175.

Barzilai, J., Cook, W., & Golany, B. (1987). Consistent weights for judgement matrices of the relative importance for alternatives. *Operations Research Letters*, 6:3, 131-134.

Belton, V., & Stewart, T.J. (2002). *Multi Criteria Decision Analysis – An Integrated Approach*. Kluwer Academic Publishers.

Checkland, P. (2001). Soft systems methodology. In J. Rosenhead, & J. Mingers (Eds.), *Rational analysis for a problematic world revisited* (pp. 61-91). (2nd ed.). John Wiley & sons.

Chen, Y., Kilgour, D.M., & Hipel, K.W. (2008). Screening in multiple criteria decision analysis. *Decision Support Systems*, 45, 278-290.

Danish Ministry of Transport (2003). Manual for Socio Economic Appraisal: Applied Methodology in the Transport Sector (in Danish). Danish Ministry of Transport, Copenhagen.

Eden, C., & Ackermann, F. (2001). SODA – the principles. In J. Rosenhead, & J. Mingers (Eds.), *Rational analysis for a problematic world revisited* (pp. 21-43). (2nd ed.). John Wiley & sons.

Friend, J. (2001). The strategic choice approach. In J. Rosenhead, & J. Mingers (Eds.), *Rational analysis for a problematic world revisited* (pp. 115-151). (2nd ed.). John Wiley & sons.

Goodwin, P., & Wright, G. (2009). *Decision Analysis for Management Judgment*. (4th ed.). Wiley.

Guitouni, A., & Martel, J-M. (1998). Tentative guidelines to help choosing an appropriate MCDA method. *European Journal of Operational Research*, 109, 501-521.

Hobbs, B.F., & Meier, P. (2000). *Energy decisions and the environment: A guide to the use of multicriteria methods*. Kluwer Academic Publishers.

Keeney, R.L. (1992). *Value focused thinking*. Harvard University Press.

Keeney, R.L., & Raiffa, H. (1993). *Decisions with Multiple Objecttives – Preferences and Value Tradeoffs*. Cambridge University Press.

Leleur, S., Petersen, N.B., & Barfod, M.B. (2007). The COSIMA Approach to Transport Decision Making – Combining Cost-benefit and Multi-criteria Analysis for Comprehensive Project Appraisal. In K-S. Kim (Ed.), *Improving Public Investment Management for Large-Scale Government Projects: Focusing on the Feasibility Studies – Lessons and Challenges* (pp. 100-122). KDI.

Leleur, S. (2008). Systemic Planning. Second edition, Polyteknisk Press.

Lootsma, F.A. (1988). Numerical scaling of human judgement in pairwise-comparison methods for fuzzy multi-criteria decision analysis. In G. Mitra (ed.), *Mathematical Models for Decision Support* (pp. 57-88). Springer-Verlag.

Lootsma, F.A. (1999). *Multi-criteria decision analysis via ratio and difference judgment*. Kluwer Academic Publishers.

Løken, E. (2007). Use of multicriteria decision analysis methods for energy planning problems. *Renewable & Sustainable Energy Reviews*, 11, 1584-1595.

Mingers, J., & Rosenhead, J. (2004). Problem structuring methods in action. *European Journal of Operational Research*, 152, 530-554.

Mustajoki, J., Hämäläinen, R. P., & Sinkko, K. (2007). Interactive computer support in decision conferencing: Two cases on off-site nuclear emergency management. *Decision Support Systems*, 42, 2247-2260.

Neves, L.P., Dias, L.C., Antunes, C.H., & Martins, A.G. (2009). Structuring an MCDA model using SSM: A case study in energy efficiency. *European Journal of Operational Reseach*, 199, 834-845.

Olson, D.L., Fliedner, G., & Currie, K. (1995). Comparison of the REMBRANDT system with analytic hierarchy process. *European Journal of Operational Research*, 82, 522-539.

Phillips, L.D. (1984). A theory of requisite decision models. *Acta Psychologica*, 56, 29-48.

Phillips, L.D. (2007). Decision conferencing. In W. Edwards, R.F. Miles, & D. von Winterfeldt (Eds.), *Advances in Decision Analysis* (pp. 375-399). Cambridge University Press.

Ramanathan, R. (1997). Stochastic decision making using multiplicative AHP. *European Journal of Operational Research*, 97, 543-549.

Rosenhead, J., & Mingers, J. (2001). A new paradigm of analysis. In J. Rosenhead, & J. Mingers (Eds.), *Rational analysis for a problematic world revisited* (pp. 1-21). (2nd ed.). John Wiley & sons.

Roy, B. & Vanderpooten, D. (1996). The European School of MCDA: Emergence, Basic Features, and Current Works. *Journal of Multi-Criteria Decision Analysis,* 5, 22-37.

Saaty, T. L. (1977). Scenarios and priorities in transport planning: Application to the Sudan. *Transport Research*, 11, 343-350.

Saaty, T. L. (2001). *Decision Making for Leaders: The Analytical Hierarchy Process for Decisions in a Complex World*. The Analytical Hierarchy Process Series.

Stewart, T.J., & Losa, F.B. (2003). Towards reconciling outranking and value measurement practice. *European Journal of Operational Research*, 145, 645-659.

van den Honert, R.C., & Lootsma, F.A. (2000). Assessing the quality of negotiated proposals using the REMBRANDT system. *European Journal of Operational Research*, 120, 162-173.

von Winterfeldt, D., & Edwards, W. (1986). *Decision Analysis and Behavioral Research*. Cambridge University Press

von Winterfeldt, D., & Edwards, W. (2007). Defining a Decision Analytic Structure. In W. Edwards, R.F. Miles, & D. von Winterfeldt (Eds.), *Advances in Decision Analysis* (pp. 81-103). Cambridge University Press.

von Winterfeldt, D., & Fasolo, B. (2009). Structuring decision problems: A case study and reflections for practitioners. *European Journal of Operational Research*, 199, 857-866.

Zopounidis, C., & Doumpos, M. (2002). Multicriteria classification and sorting methods: A literature review. *European Journal of Operational Research*, 138, 229-246.

Paper 5

Scaling transformation in the REMBRANDT technique: a sensitivity examination of the progression factors

Michael Bruhn Barfod
Department of Transport, Technical University of Denmark

Steen Leleur
Department of Transport, Technical University of Denmark

Paper submitted for publication in *International Journal of Information Technology & Decision Making,* November 2011, the paper is currently under review – ISI-indexed

Keywords: Multi-criteria decision analysis, decision support systems, REMBRANDT

Abstract

This paper examines a decision support system (DSS) for the appraisal of complex decision problems using multi-criteria decision analysis (MCDA). The DSS makes use of a structured hierarchical approach featuring the multiplicative AHP also known as the REMBRANDT technique. The paper addresses the influence of the progression factor used when transforming the decision-makers verbal responses from a semantic to a geometric scale using the technique. Conventionally, the progression factor 2 is used for calculating scores of alternatives and the square root of 2 for calculation of criteria weights. Tests are conducted on the magnitude of these progression factors in order to examine the sensitivity towards the final outcome of an analysis. For illustration a case study dealing with the appraisal task of a large transport infrastructure project between Denmark and Sweden is presented. Finally, conclusions are drawn and perspectives are set out in the context of the proposed DSS and its use for strategic decision making.

1. Introduction

This paper examines a decision support system (DSS) for the appraisal of complex transport infrastructure decision problems using multi-criteria decision analysis (MCDA). The DSS makes use of a structured hierarchical approach featuring the multiplicative AHP also known as the REMBRANDT technique. The technique is a further development of the original AHP and it proposes to overcome three issues regarding the theory behind AHP namely by using direct rating on a geometric scale, the geometric mean method, and aggregation of scores by the product of alternative relative scores weighted by the power of weights obtained from analysis of the hierarchical elements above the alternatives. The aim of this paper is mainly to address the first issue regarding the direct rating on a geometric scale.

More specifically, the paper addresses the influence of the progression factors used when transforming the decision-makers' verbal responses from the semantic to the geometric scale. The REMBRANDT technique uses the progression factor 2 for calculating scores of alternatives and the square root of 2 for calculation of criteria weights, where the reason behind a lower progression factor for criteria may link to implicit trade-off considerations being more deliberate with criteria than is the case with scoring of alternatives. Tests will be conducted on the magnitude of the progression factors in order to examine the sensitivity towards the final outcome of an analysis.

For illustration of the DSS and the sensitivity calculations a case study dealing with the appraisal task of a large transport infrastructure project is presented. The scope of the case study is to identify the most attractive alternative for a new bridge or tunnel connection between the cities of Elsinore (Helsingor) in Denmark and Helsingborg in Sweden, which is supposed to take over both person and freight transport from the existing ferries and relieve the existing fixed link between Copenhagen in Denmark and Malmo in Sweden. The appraisal will make use of previously conducted cost-benefit calculations and descriptions of strategic issues. Finally, conclusions are drawn and research issues defining future work are set out in the context of the proposed DSS and its use for strategic decision making.

This paper is disposed as follows: After this introduction Section 2 introduces the case study used for illuminating the test calculations. Section 3 contains a description of the the REMBRANDT technique (the multiplicative AHP) and presents more closely the scaling issues of the progression factors used for transformations to the geometric scale. In Section 4 the REMBRANDT technique is applied to the case study and sensitivity calculations are made based on the progression factors. The results of these calculations are subsequently discussed in Section 5 before Section 6 concludes and outlines perspectives for future work within the research area.

2. The case study

The Oresund fixed link connecting the greater area of Copenhagen with Malmo in Sweden opened in July 2000. Today, eleven years later, the railway line of the link is close to its capacity limit resulting in delays and discomfort for the travellers. The case of this paper concerns a new complementary fixed link between Denmark and Sweden between the cities of Elsinore (Helsingor) and Helsingborg. Regionally, the proposed connection is expected to create a substantial increase in trade, education and work related benefits. Ultimately it is expected that a fixed link with increased commuter traffic across the border will result in a common labour and residence market. In addition, the recent decision to construct the Fehmarn Belt fixed link between Denmark and Germany will increase the number of travellers from central Europe through Denmark to Sweden, Norway and Finland. This means additional traffic to cross the Oresund (Larsen and Skougaard, 2010).

The case is normally referred to as the HH-connection, see Figure 1, and has been examined since the 1980s where the first alignment proposals were suggested. The opening of the Oresund fixed link between Copenhagen and Malmo, however, postponed the HH-connection but now its planning is recommenced. In Figure 1 the proposed new fixed link is shown located approximately 50 km north of the existing fixed link across Oresund.

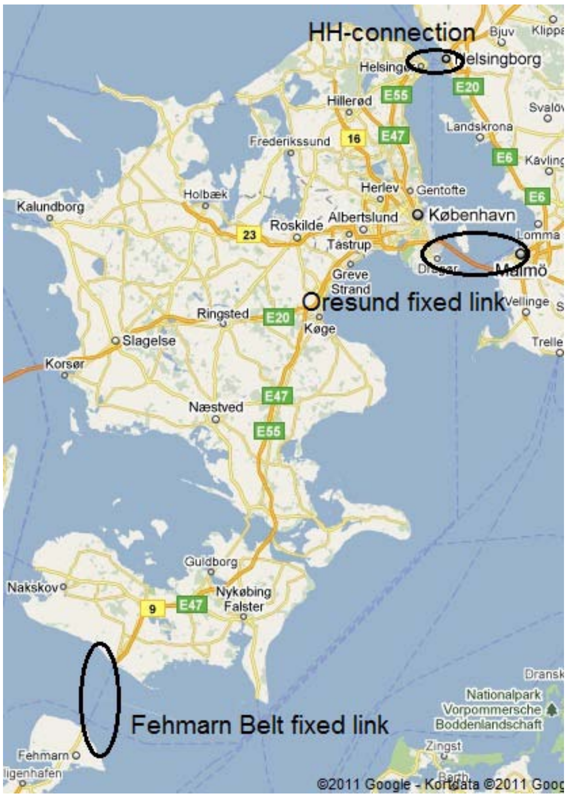


Figure 1. The proposed new fixed link (the HH-connection), the Oresund fixed link and the forthcoming Fehmarn belt fixed link (from Google maps)

The current situation with ferry service is referred to as the basis scenario where the proposed alternatives will substitute the ferries with a fixed link with four alternatives being considered, see Table 1.

Table 1. The four proposed alternatives for the HH-connection with construction costs in million DKK (Larsen and Skovgaard, 2010)

HH-connection	Description	Cost (mDKK)
Alternative 1 (A1)	Tunnel for rail (2 tracks) person traffic only	7,700
Alternative 2 (A2)	Tunnel for rail (1 track) goods traffic only	5,500
Alternative 3 (A3)	Bridge for road and rail (2x2 lanes and 2 tracks)	11,500
Alternative 4 (A4)	Bridge for road (2x2 lanes)	6,000

Based on the Danish manual for socio-economic assessment (DMT, 2003) the benefit-cost rates (BCR) shown in Table 2 have been determined applying transport modelling for road and rail (Salling et al., 2010).

The reference scenario forecasting (RSF) technique (Salling and Leleur, accepted) is applied to the cost-benefit analysis in order to produce certainty graphs describing the alternatives' probability for being economic feasible if uncertainties are introduced to the assessment. The calculations are carried out by applying Monte Carlo simulation to the cost-benefit analysis using estimated RSF distributions (Erlang and Beta-Pert distributions for construction costs and time savings respectively). Based on this it is possible to calculate certainty values (CV), see Table 2, which are describing the probability for the alternatives being feasible.

Table 2. BCRs for the four alternatives

Alternatives	BCR	CV
A1	1.51	97%
A2	0.18	10%
A3	2.72	123%
A4	3.09	157%

It should be noted that an alternative will obtain a CV > 100% if the simulation implies BCRs > 1 in all cases. For CVs > 100% the certainty 'distance' above 100% is added to the value. Where the conventional BCR gives a deterministic point estimate of the feasibility, the CVs give a probability based interval estimate of how the two most important uncertainty factors can affect such a point estimate.

Due to the high influence on the further development of the Oresund region a wider set of decision criteria have been identified in addition to the CVs to lay the foundation for a comprehensive assessment of the four alternatives. Table 3 depicts the full criteria-set (Larsen and Skovgaard, 2010).

Table 3. Assessment criteria

Criterion	Definition
C1	Impact on regional economics
C2	Impact on ecology in sound
C3	Impact on transport network and accessibility
C4	Impact on towns
C5	Robustness of feasibility (CVs)

The alternatives are in Section 4 assessed under the criteria using the multiplicative AHP also known as the REMBRANDT technique, which is described in the following Section 3.

3. The REMBRANDT technique (multiplicative AHP)

The original AHP by Saaty (1977; 2009) has been criticised for various reasons: 1) for the fundamental scale to quantify human judgments; 2) as it estimates the impacts scores of the alternatives by the Perron-Frobenius eigenvector; and 3) as it calculates the final scores of the alternatives using the arithmetic-mean aggregation rule. These controversial issues are well-known and not new. Already Zahedi (1986) signalised that the criticism of the AHP concentrated on the estimation of the impact scores, but that no major controversy existed concerning the aggregation step. Criticism of the fundamental scale was not mentioned by Zahedi, but Belton (1986) brought forward several arguments against the scale and the aggregation rule. Later also Stewart (1992) discussed the above issues and warned that the AHP, despite its widespread popularity, should be used with considerable caution. More recently Korhonen and Topdagi (2003) have also brought forward arguments regarding the inappropriateness of the ratio scale in specific decision situations. Barzilai et al. (1987), Barzilai and Golani (1991) and Barzilai (1992) observed that the AHP, since it is initially based upon ratio information, should be converted into a variant with a multiplicative structure.

A multiplicative version of the original AHP is available in form of the so-called REMBRANDT (Ratio Estimations in Magnitudes or deci-Bells to Rate Alternatives which are Non-DominaTed technique), see Lootsma (1992), Olson et al. (1995) and Ramanathan (1997). As for the original AHP the REMBRANDT technique makes use of a structured hierarchical approach based on the principle that decision-makers make pair wise comparisons between alternatives to determine subjective impacts under each criterion in the assessment and between criteria in order to determine their relative importance. Finally, aggregating the results leads to a final score for each project, which allows a subjective rank ordering of the projects.

The systematic pair wise comparison approach is one of the cornerstones of the REMBRANDT technique (Lootsma, 1992). REMBRANDT makes use of a procedure for direct rating which requires the decision-makers to consider all possible pairs of alternatives with respect to each criterion in turn in order to determine which one of the projects in the pair is preferred and to specify the strength of preference according to a semantic scale (associated a numeric 0-8 scale). The approach is as mentioned a multiplicative development of the AHP and it proposes to overcome the three issues regarding the theory behind AHP.

First, the direct rating in REMBRANDT is on a geometric scale (Lootsma, 1992) which replaces Saaty's 1 – 9 original scale. Second, the eigenvector method originally used in AHP is replaced by the geometric mean method, which avoids potential rank reversal (Barzilai et al., 1987). Third, the aggregation of scores by arithmetic mean is replaced by the product of alternative relative scores weighted by the power of weights obtained from analysis of the hierarchical elements above the alternatives (Olson, 1996).

In the use of the REMBRANDT technique in this paper it is assumed that the ratifying group consists of g decision-makers ($g \geq 1$), and that at any stage of the process there are n alternatives ($n \geq 1$) under consideration. At the first evaluation level of the analysis, each pair of alternatives A_j and A_k is presented to the decision-makers under a specific criterion. The decision-makers are then asked to express their graded comparative judgment about them. That is, the decision-makers express their indifference between the two, or a weak, definite, strong or very strong preference for one project over the other. Thus, at this stage the decision-makers are asked to make as standard $n(n-1)/2$ pair wise comparisons. Indeed only *(n-1)* properly chosen comparisons would be sufficient, for which reason the standard leads to much more information being collected than actually needed (Zahir, 2006). Such redundancy, however, is usually beneficial as it enables a smoothing of the results of the analysis. Incomplete pair wise comparisons in a group of decision-makers are handled in a general way by using REMBRANDT, see (Lootsma, 1999) for details; the case of complete pair wise comparisons by each and every one of the decision-makers is a special case. In this context it is assumed that alternative A_j and A_k have the same subjective values V_j and V_k for all decision-makers in a group. Using the REMBRANDT technique the group's agreed upon judgment about the pair A_j and A_k is taken to be an estimate of the preference ratio V_j/V_k.

The decision-makers' pair wise comparative judgment of A_j versus A_k is captured on a category scale to frame the range of possible verbal responses. This is converted into an integer-valued gradation index δ_{jk} according to the REMBRANDT scale in Table 4. The number of categories is rather small as human beings' linguistic capacity to describe the categories unambiguously in verbal terms is limited (Ibid.).

Table 4. The REMBRANDT scale (Lootsma, 1999)

Comparative judgment	**Gradation index δ_{jk}**
Very strong preference for A_k over A_j	-8
Strong preference for A_k over A_j	-6
Definite preference for A_k over A_j	-4
Weak preference for A_k over A_j	-2
Indifference	0
Weak preference for A_j over A_k	+2
Definite preference for A_j over A_k	+4
Strong preference for A_j over A_k	+6
Very strong preference for A_j over A_k	+8

Intermediate integer values can be assigned to δ_{jk} to express a hesitation between two adjacent categories. The gradation index δ_{jk} can be converted into a value on a geometric scale, characterised by a scale parameter $\gamma = ln(1+\varepsilon)$, where $1+\varepsilon$ is the progression factor. Thus

$$r_{jk} = exp\ (\gamma\delta_{jk}), \qquad j, k = 1, \ldots, n$$

is defined to be the numeric estimate of the preference ratio V_j/V_k. Although there is no unique scale of human judgment, a plausible value of γ is *ln(2)* implying a geometric scale with the progression factor 2 (Lootsma, 1992).

There are five major, linguistically distinct categories in Table 3: indifference, weak, definite, strong and very strong. Moreover, there are four so-called threshold categories between them which can be used if the decision-makers are in-between the neighbouring qualifications. Lootsma (1999) shows that human beings follow the same pattern in many unrelated areas when they categorise an interval, e.g. certain ranges on the time axis and sound and light intensities. Normally three to five major categories are introduced and the progression factor $exp(2\gamma) = (1 + \varepsilon)^2$ is roughly 4, see Lootsma (1992, 1999). By the interpolation of threshold categories a more refined subdivision of the given interval is obtained. In that case there are six to nine categories and the progression factor $exp(\gamma) = (1 + \varepsilon)^2$ is roughly 2 $(\gamma = ln2 \sim 0.7)$, which defines what Lootsma (1993) calls the natural REMBRANDT scale. In addition, Lootsma (1993) suggest that sensitivity analysis should be carried out with a short $(\gamma = 0.5)$ and a long $(\gamma = 1.0)$ geometric scale in the neighbourhood of the natural scale.

When determining criteria weights Lootsma (1999) finds the progression factor to be √2. The reason behind a lower progression factor may link to implicit trade-off consideration being more deliberate with criteria than is the case with scoring of alternatives.

In the psychophysical literature the issue of how human beings judge the relationship between two stimuli in a pair wise comparison on one single dimension was first treated 50 years ago by Torgerson (1961). Torgerson observed that human beings perceive only one quantitative relation, but they estimate differences in subjective stimulus values when they are requested to express their judgement on a category scale with arithmetic progression (equidistant echelons) and they estimate ratios of subjective stimulus values when the proposed scale is geometric. Thus they interpret the relationship as it is required in the experiment. Which of the two interpretations is correct cannot empirically be decided, as they are alternative ways of saying the same thing.

This observation is easy to understand if it is assumed that the subjective stimulus values are not identically used in the two types of experiments. In the ratio experiment with a geometric scale human beings judge the ratio of two stimulus values. In the difference experiment with an arithmetic scale they do not judge the ratio itself but its order of magnitude, which is essentially a logarithm of the ratio (Lootsma, 1993).

Veit (1978) and Birnbaum (1982) confirmed Torgerson's observation that pair wise comparative judgment of two stimuli uses one operation only in both types of experiments. Moreover, if subtraction is assumed to be the underlying operation, then the ratio judgment is exponentially related to difference judgment.

4. Applying the REMBRANDT technique on the case study

To demonstrate the approach the four alternatives, A1, A2, A3 and A4, are compared in a pair wise way under the five criteria outlined in Table 3. A decision conference approach as described in Barfod and Leleur (2009) and Phillips (2007) was used for the purpose of deriving preferences from the decision-makers involved in the decision process. Table 5 – 9 shows the δ_{jk} matrices based on the decision-makers judgments as well as the transformed matrices and the row-wise geometric means.

Table 5. REMBRANDT calculations for C1: Impact on regional economics

	Pair wise comparisons (δ_{jk})				Transformations (γ = 0.7)				Geo.mean
	A1	A2	A3	A4	A1	A2	A3	A4	Score
A1	0	4	-4	-2	1	16	0.0625	0.25	0.71
A2	-4	0	-8	-6	0.0625	1	0.0039	0.0156	0.04
A3	4	8	0	2	16	256	1	4	11.31
A4	2	6	-2	0	4	64	0.25	1	2.83

Table 6. REMBRANDT calculations for C2: Impact on ecology in sound

	Pair wise comparisons (δ_{jk})				Transformations (γ = 0.7)				Geo.mean
	A1	A2	A3	A4	A1	A2	A3	A4	Score
A1	0	0	4	4	1	1	16	16	4.00
A2	0	0	4	4	1	1	16	16	4.00
A3	-4	-4	0	0	0.0625	0.0625	1	1	0.25
A4	-4	-4	0	0	0.0625	0.0625	1	1	0.25

Table 7. REMBRANDT calculations for C3: Impact on transport network and accessibility

	Pair wise comparisons (δ_{jk})				Transformations (γ = 0.7)				Geo.mean
	A1	A2	A3	A4	A1	A2	A3	A4	Score
A1	0	6	-4	-2	1	64	0.0625	0,25	1.00
A2	-6	0	-8	-6	0.0156	1	0.0039	0.0156	0.03
A3	4	8	0	4	16	256	1	16	16.00
A4	2	6	-4	0	4	64	0.0625	1	2.00

Table 8. REMBRANDT calculations for C4: Impact on towns

	Pair wise comparisons (δ_{jk})				Transformations (γ = 0.7)				Geo.mean
	A1	A2	A3	A4	A1	A2	A3	A4	Score
A1	0	-2	6	4	1	0.25	64	16	4.00
A2	2	0	8	6	4	1	256	64	16.00
A3	-6	-8	0	-2	0.0156	0.0039	1	0.25	0.06
A4	-4	-6	2	0	0.0625	0.0156	4	1	0.25

Table 9. REMBRANDT calculations for C5: Robustness of feasibility

Pair wise comparisons (δ_{jk})					**Transformations (γ = 0.7)**				**Geo.mean**
	A1	A2	A3	A4	A1	A2	A3	A4	Score
A1	0	6	-4	-6	1	64	0.0625	0.0156	0.50
A2	-6	0	-6	-8	0.0156	1	0.0156	0.0039	0.03
A3	4	6	0	-2	16	64	1	0.25	4.00
A4	6	8	2	0	64	256	4	1	16.00

This is followed by pair wise comparisons of the five criteria in Table 10.

Table 10. REMBRANDT calculations for criteria weights

Pair wise comparisons						**Transformations (γ = 0.35)**					**Geo. mean**	
	C1	C2	C3	C4	C5	C1	C2	C3	C4	C5	Score	Norm.
C1	0	4	-2	2	-2	1	4	0.5	2	0.5	1.15	0.19
C2	-4	0	-4	-3	-4	0.25	1	0.25	0.3536	0.25	0.35	0.06
C3	2	4	0	3	-2	2	4	1	2.8284	0,5	1.62	0.27
C4	-2	3	-3	0	-3	0.5	2.8284	0.3536	1	0.3536	0.71	0.12
C5	2	4	2	3	0	2	4	2	2.8284	1	2.14	0.36

The final scores for the alternatives are calculated using the multiplicative model and normalised:

A1: $0.71^{0.19} \cdot 4.00^{0.06} \cdot 1.00^{0.27} \cdot 4.00^{0.12} \cdot 0.50^{0.36}$ = 0.93 ~ 0.12

A2: $0.04^{0.19} \cdot 4.00^{0.06} \cdot 0.03^{0.27} \cdot 16.00^{0.12} \cdot 0.03^{0.36}$ = 0.09 ~ 0.01

A3: $11.31^{0.19} \cdot 0.25^{0.06} \cdot 16.00^{0.27} \cdot 0.06^{0.12} \cdot 4.00^{0.36}$ = 3.69 ~ 0.47

A4: $2.83^{0.19} \cdot 0.25^{0.06} \cdot 2.00^{0.27} \cdot 0.25^{0.12} \cdot 16.00^{0.36}$ = 3.12 ~ 0.40

By normalising the REMBRANDT scores above we arrive at the score-set: A1 = 0.12; A2 = 0.01; A3 = 0.47; A4 = 0.40. If the same verbal responses instead had been processed using the AHP technique the scores would have been: A1 = 0.15; A2 = 0.11; A3 = 0.38; A4 = 0.36, which is relatively close to the REMBRANDT scores, see Table 11. The basic observation here is that using the REMBRANDT technique the best performing alternatives seems to get an advantage due to the longer scale resulting in greater interval distance.

Table 11. Scores for the alternatives calculated using AHP and REMBRANDT applying different progression factors for the alternatives

Alternative	AHP	REMBRANDT		
		γ = 0.5	γ = 0.7	γ = 1.0
A1	0.15	0.19	0.12	0.04
A2	0.11	0.04	0.01	0.01
A3	0.38	0.41	0.47	0.52
A4	0.36	0.36	0.40	0.43

As mentioned Lootsma (1993) suggests to conduct sensitivity analysis with γ = 0.5 (progression factor on 1.6) and γ = 1.0 (progression factor on 2.7) to test the robustness of the results. The outcome of this is also shown in Table 11.

To examine the sensitivity of the progression factor in a wider interval tests have been conducted varying the factor from 1 to 5 (γ values between 0.0 and 1.6), see Figure 2. From this the important result that the rank order of the alternatives does not depend on the scale parameter γ can be derived.

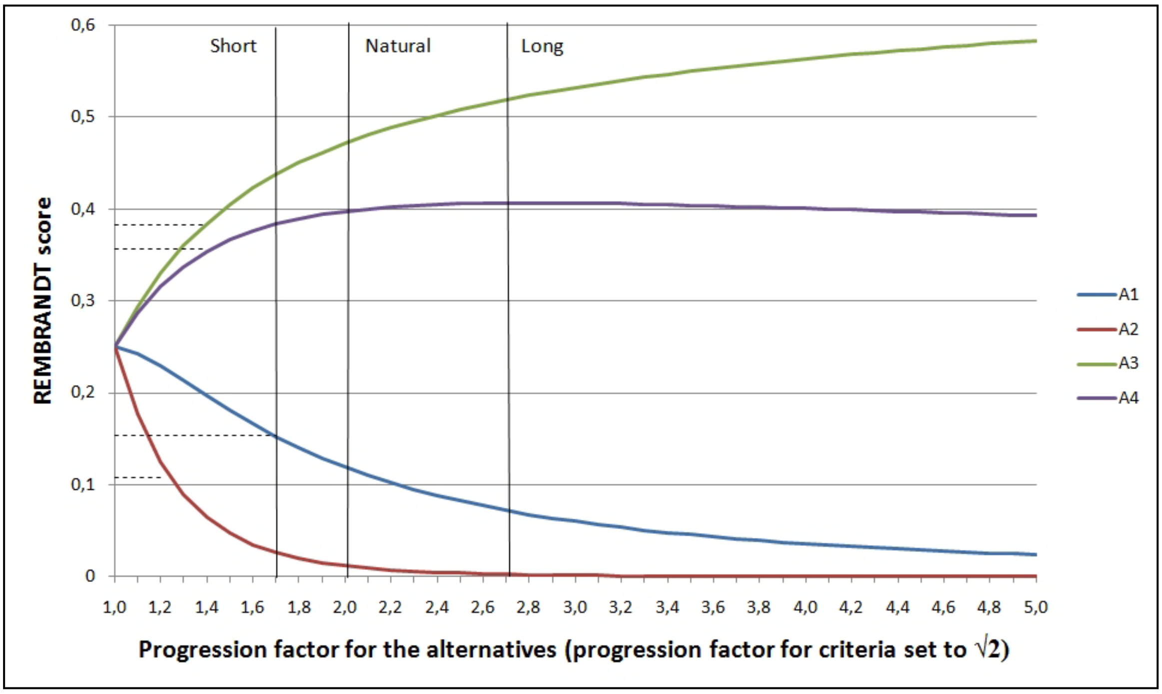


Figure 2. REMBRANDT scores at varying values of the progression factor for the alternatives. The vertical lines indicate the short, the natural and the long REMBRANDT scales proposed by Lootsma (1993). The horizontal dashed lines indicate the corresponding AHP scores (see Table 11).

It should be noted that the progression factor for the criteria weights is not varied in this sensitivity test as there is only proposed one particular geometric scale for this (γ = 0.35), not a variety of scales to quantify the gradations of comparative judgments. In practice the difference between pair wise comparisons at the first and second evaluation levels is small. There are conceptual differences between the evaluation levels, and this implies that the numerical values of the quantifiers are level dependent (Lootsma, 1993).

5. Discussion of results

Observing Figure 2 it seems reasonable that the progression factor should not exceed 2.7, as indicated by Lootsma (1993), by much. Higher factors will increase then span between the worst performing and the best performing alternatives more than seems appropriate. Similarly, progression factors close to 1 do not seem appropriate as the segregation between both alternatives and criteria will be very narrow leading to almost identical scores. In fact, the interval proposed by Lootsma (1993) seems intuitively appropriate with a short, a natural and a long scale. The calculations clearly illustrate that the ratio of two final scores is scale dependent, even under conditions which guarantee that it is not affected by the addition or deletion of alternatives.

As mentioned, there is only proposed one geometric scale for the criteria weights. However, as this scale with the progression factor√2 seems to be a result of the mathematics behind the method (Olson et al., 1995) it could also be interesting to conduct a sensitivity analysis on this. In Figure 3 the progression factor for the alternative is fixed to 2, while the progression factor for the criteria is varied in an interval from 1 to 2.3 (as no changes in rankings takes place after this point.

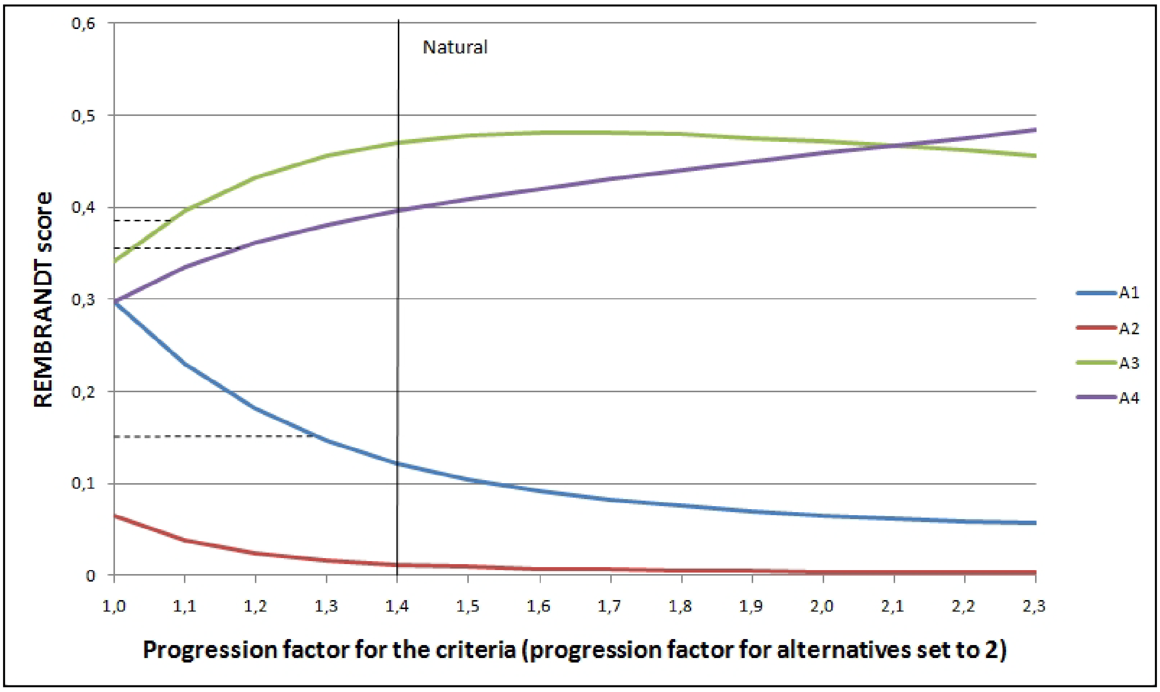


Figure 3. REMBRANDT scores at varying values of the progression factor for the criteria. The vertical line indicates the natural REMBRANDT scale with the progression factor√2. The progression factor for the alternatives is fixed to 2. The horizontal dashed lines indicate the corresponding AHP scores (see Table 11).

The results in Figure 3 clearly show that the rank order of the alternatives does depend on the scale parameter γ for the criteria. This makes good sense and it can be noted that the rank reversal which takes place between A3 and A4 at a progression factor on approximately 2.1 is caused by the fact that the weight for C5 (the criterion with the highest weight) becomes very dominant at high progression factors. Hence, the remaining criteria will move towards exclusion from the analysis, and the alternative which scores the best under C5 (A4 in this case) will be the most attractive. A3 is in this respect only the second highest scoring alternative under C5 (see Table 9). Thus, a progression factor for the criteria that exceeds √2 will not be appropriate in practical use.

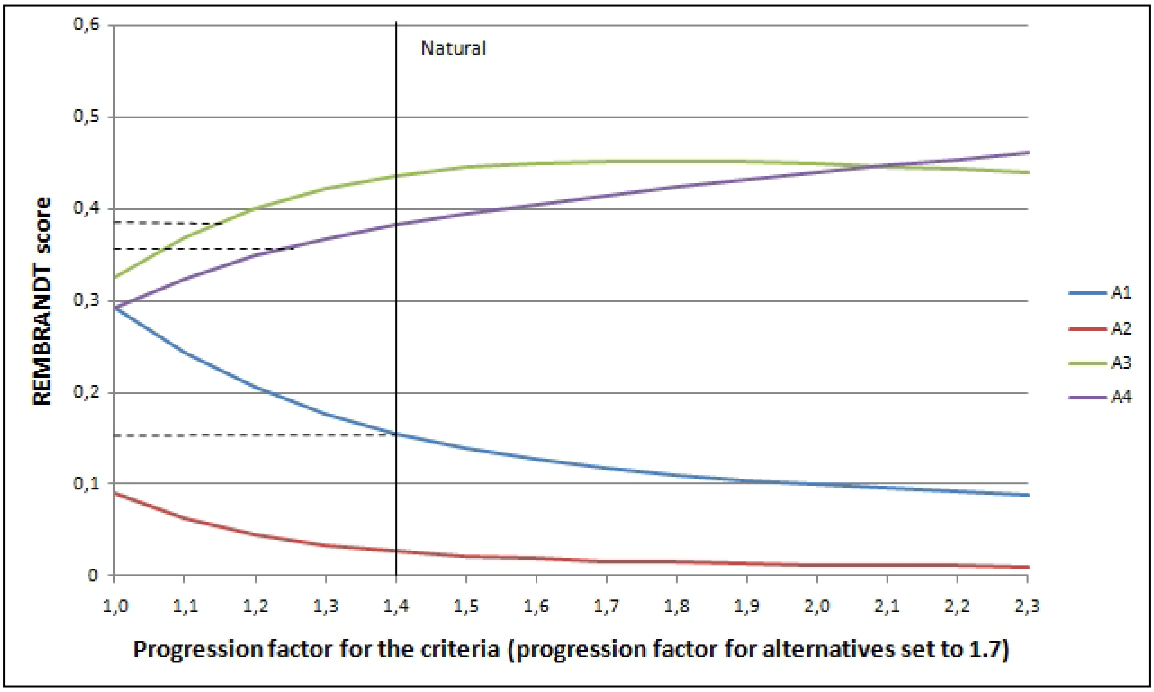


Figure 4. REMBRANDT scores at varying values of the progression factor for the criteria. The vertical line indicates the natural REMBRANDT scale with the progression factor√2. The progres sion factor for the alternatives is now fixed to 1.7 (the short scale). The horizontal dashed lines indicate the corresponding AHP scores (see Table 11).

The results imply that a modified version of REMBRANDT could make use of the progression factors 1.7 and 1.3 for alternatives and criteria respectively in order to obtain results closer in line with the results from the original AHP.

In order to test this argument another case example, which concerns four alternatives for a railway line assessed under eight criteria, is examined, see Table 12. The case is described in details in Barfod et al. (2011a).

Table 12. Scores for the alternatives in Barfod et al. (2011a) calculated using AHP and REMBRANDT applying different progression factors.

Alternative	AHP	REMBRANDT			
		A-prog. 2 C-prog. √2	A-prog. 1.7 C-prog. √2	A-prog. 1.7 C-prog. 1.3	A-prog. 1.7 C-prog. 1.6
R	0.26	0.17	0.19	0.16	0.24
BS	0.29	0.39	0.36	0.38	0.34
BL	0.28	0.33	0.31	0.35	0.28
G	0.17	0.11	0.14	0.14	0.14

The results in Table 12 are not in support of the previous results that argued for a lowering of the criteria progression factor. This might very well be due to a higher number of criteria in the assessment. However, the results still imply that a lowering of the alternative progression factor to 1.7 seems reasonable.

One more case study featuring four alternatives for a fixed link assessed under four criteria is examined in Table 13 to test the robustness of the arguments. The case is described in details in Barfod et al. (2011b).

Table 13. Scores for the alternatives in Barfod et al. (2011b) calculated using AHP and REMBRANDT applying different progression factors

Alternative	AHP	REMBRANDT			
		A-prog. 2 C-prog. √2	A-prog. 1.7 C-prog. √2	A-prog. 1.7 C-prog. 1.3	A-prog. 1.7 C-prog. 1.6
HL	0.44	0.69	0.68	0.60	0.76
ST	0.23	0.15	0.16	0.18	0.14
LT	0.18	0.07	0.06	0.10	0.03
UP	0.15	0.09	0.09	0.12	0.06

The REMBRANDT results in Table 13 differ much more from the original AHP results than has been the case with the two previous examined studies. The difference concerns both the size of the scores and the implied ranking, see the ranking of the alternatives on level three and four. However, the results are in line with the previous results in the sense that a lower progression factor for the alternatives seems reasonable.

Finally, a forth case study is examined. The case, which is described in details in Barfod (2012), concerns five alternatives for bike projects assessed under five criteria. Table 14 depicts the results derived.

Table 14. Scores for the alternatives in Barfod (2012) calculated using AHP and REMBRANDT applying different progression factors

Alternative	AHP	REMBRANDT			
		A-prog. 2 C-prog. √2	A-prog. 1.7 C-prog. √2	A-prog. 1.7 C-prog. 1.3	A-prog. 1.7 C-prog. 1.6
ID1	0.06	0.04	0.06	0.05	0.05
ID2	0.16	0.14	0.16	0.17	0.16
ID3	0.19	0.19	0.20	0.20	0.20
ID4	0.38	0.45	0.39	0.39	0.40
ID5	0.21	0.18	0.19	0.19	0.19

The results in Table 14 show almost insignificant sensitivity towards the size of the progression factors. This may be caused by the fact that the ratifying group doing the comparisons consisted of government officers and that the decision problem was of a politically sensitive nature. For this reason the group tended to apply the semantic scale with highest precaution using only the lower values on the scale. Hence, the segregation between the alternatives is low no matter which of the approaches, AHP or REMBRANDT, is applied. However, observing Table 14 a progression factor for the alternatives on 1.7 seems most reasonable.

Overall, the results of the different case studies show that the REMBRANDT technique moderates the valuation of “extreme” versus “balanced” alternatives. In the additive AHP it may be mathematically impossible for “middle of the road” alternatives to achieve the highest overall ratings. This makes little sense from a practical viewpoint. The multiplicative version ensures due consideration of “middle of the road” non-dominated alternatives when these are assessed with alternatives that are extremely attractive with respect to some criteria and extremely un-attractive with regard to other ones. Thus, in addition to mathematical considerations (ratio scale property) there appears to be important behavioural motivation for using the REMBRANDT technique.

In practice “middle of the road” alternatives may very well be the most preferred ones; however, the flexibility of the REMBRANDT technique with varying γ-values appears to offer the decision-makers an attractive modelling framework. Previous studies of the REMBRANDT technique, such as Lootsma (1993) and Stam and Silva (2003), have implicitly assumed that the preference ratings are the geometric means of pair wise comparisons exhibiting constant returns to scale. This is a requirement which may be reasonable in many decision problems (Stam and Silva, 2003). Nevertheless, allowing for a flexible value of γ allows for a meaningful analysis of situations where increasing or decreasing returns to scale are appropriate.

6. Conclusion

This paper has examined a DSS for the appraisal of complex decision problems using MCDA. More specifically the multiplicative version of AHP, namely the REMBRANDT technique, has been examined with regard to the issue of the progression factor when transforming decision-makers verbal responses to the geometric scale. AHP was first introduced by Saaty (1977) based on an additive value function model. Several improvements of the technique have been made over the years, e.g. the introduction of the geometric mean method by Barzilai et al. (1987). With the REMBRANDT technique (based on a multiplicative utility function) by Lootsma (1992) a serious off-spring alternative to the original AHP was introduced.

Both the original AHP and the REMBRANDT techniques can be considered as effective DSSs for group decision making. The additive AHP allows a multi-level hierarchy; however, this is hardly an advantage as decision-makers tend to insist on a one-level hierarchy as this seems more intuitive. Moreover, as noted in Section 5, the final scores calculated by the two versions of AHP do not strongly diverge. However, the aggregation rule used by REMBRANDT seems appropriate as it fits the exponential form of the input given by the decision-makers.

In the ease of use the two versions are very similar as they need the same type of input and provide the same type of output. The original additive AHP has one scale only and ignores scale dependence, whereas the REMBRANDT technique, based on a one parametric class of geometric scales, yields a scale-independent rank order of the final scores and avoids rank reversal in some notorious cases where this phenomenon is not expected to occur. Seen from a theoretical viewpoint the geometric least squares method of REMBRANDT is preferable. However, in practice it does not seem to make much of a difference which method is selected.

Based on the case studies it can be recommended to conduct sensitivity analysis applying different progression factors on both the alternatives scores level and the criteria weights level. The rank order of the alternatives does not depend on the scale parameter γ when this is changed for the alternatives score level. However, it can be concluded that the scale parameter should not exceed the long scale (a progression factor on 2.7) by much as the span between the scores becomes inappropriately large. As opposed to this the rank order of alternatives are very sensitive towards changes in the progression factor on the criteria weight level. Therefore it can be recommended that a progression factor on 1.7 (the short scale) can be applied at the alternatives score level while the criteria weight level should continue to make use of a progression factor on $\sqrt{2}$ if it is desirable to arrive at results closer in line with the original AHP.

Future research within this area should concentrate on studying the further properties of the REMBRANDT technique seen from both a theoretical and empirical point of view. The attractiveness of the multiplicative version in practice compared to the additive version will in some cases depend on the decision problem. Nevertheless, the multiplicative version with variable γ-values is consistent with well-grounded postulates of human decision making which makes it attractive to apply when approaching complex appraisal problems by using multiple pair wise comparisons.

References

Barfod, M. B. and Leleur, S. (2009). Customised DSS and decision conferences. Paper in the proceedings of the 13th Euro Working Group on Transportation (EWGT) meeting, University of Padua, Italy.

Barfod, M. B., Jensen, A. V. and Leleur, S. (2011a). Examination of Decision Support Systems for Composite CBA & MCA Assessments of Transport Infrastructure Projects. In: Shi, Y., Wang, S., Kou, G. and J. Wallenius (Eds.). *New State of MCDM in the 21st Century*. Lecture Notes in Economics and Mathematical Systems 648, pp. 167-176, Springer.

Barfod, M. B., Salling, K. B. and Leleur, S. (2011b). Composite Decision Support by Combining Cost-benefit and Multi-criteria Decision Analysis. *Decision Support Systems* 51, pp. 167-175, Elsevier.

Barfod, M. B. (2012). An MCDA approach for the selection of bike projects based on structuring and appraising activities. *European Journal of Operational Research* 218, pp. 810-818, Elsevier.

Barzilai, J., Cook, W. and Golany, B. (1987). Consistent weights for judgement matrices of the relative importance for alternatives. *Operations Research Letters* 6:3, pp. 131-134, Elsevier.

Barzilai, J. (1992). On the use of the eigenvector in the AHP. Paper in the proceedings of the International Conference on MCDM, Tapei, pp. 291-300.

Barzilai, J. and Golany, B. (1991). An axiomatic framework for aggregating weights and weight-ratio matrices. Paper in the Proceedings of the second International Symposium on AHP, Pittsburgh, pp. 59-70.

Belton, V. (1986). A comparison of analytic hierarchy process and a simple multi-attribute value function. *European Journal of Operational Research* 26, pp. 7-21, Elsevier.

Birnbaum, M. H. (1982). Controversies in psychological measurement. In: Hillsdale, B. W (Ed.). *Social attitudes and psyco-physical measurement*. pp. 401-485.

Danish Ministry of Transport (2003). Manual for socio-economic analysis – applied methodology in the transport sector. Report from the Danish Ministry of Transport, Copenhagen (in Danish).

Korhonen, P. and Topdagi, H. (2003). Performance of the AHP in comparison of gains and losses. *Mathematical and Computer Modelling* 37, pp. 757-766, Elsevier.

Larsen, L. A. and Skovgaard, B. Z. (2010). Appraisal of alternatives concerning a fixed link between Elsinore and Helsingborg. Master thesis, Department of Transport, Technical University of Denmark (in Danish).

Lootsma, F. A. (1992). The REMBRANDT system for multi-criteria decision analysis via pair wise comparisons or direct rating. Report 92-05, Faculteit der Technische Wiskunde en Informatica, Delft University of Technology, Netherlands.

Lootsma, F. A. (1993). Scale sensitivity in the multiplicative AHP and SMART. *Journal of Multi-Criteria Decision Analysis* 2, pp. 87-110, Wiley.

Lootsma, F. A. (1999). *Multi-criteria decision analysis via ratio and difference judgement.* Kluwer Academic Publishers, Dordrecht, Netherlands.

Olson, D. L., Fliedner, G. and Currie, K. (1995). Comparison of the REMBRANDT system with analytic hierarchy process. *European Journal of Operational Research* 82, pp. 522-539, Elsevier.

Olson, D. L. (1996). *Decision Aids for Selection Problems.* Springer-Verlag, New York, United States of America.

Phillips, L. D. (2007). Decision conferencing. In: Edwards, W., Miles, R. F. and von Winterfeldt, D. (Eds.). *Advances in decision analysis – from foundations to applications* (pp. 375-400). Cambridge University Press, New York, United States of America.

Ramanathan, R. (1997). Stochastic decision making using multiplicative AHP. *European Journal of Operational Research* 97, pp. 543-549, Elsevier.

Saaty, T. L. (1977). Scenarios and priorities in transport planning: Application to the Sudan. *Transport Research* 11, pp. 343-350, Pergamon Press.

Saaty, T. L. and Sagir, M. (2009). Extending the measurement of tangibles to intangibles. *International Journal of Information Technology & Decision Making* 8, pp. 7-27, World Scientific.

Salling, K. B., Leleur, S. and Skougaard, B. Z. (2010). Reference Scenario Forecasting: A new approach to Transport Project Assessment. Paper in the proceedings of the 10th World Conference on Transport Research (WCTR), Lisbon.

Salling, K. B. and Leleur, S. (accepted). Modelling of Transport Project Uncertainties: Risk Assessment and Scenario Analysis. Accepted for publication in *European Journal of Transport Infrastructure Research,* Open Access, TUDelft.

Stam, A. and Silva, A. P. D. (2003). On multiplicative priority rating methods for the AHP. *European Journal of Operational Research* 145, pp. 92-108, Elsevier.

Stewart, T. J. (1992). A critical survey on the status of multi-criteria decision making theory and practice. *Omega* 20, pp. 569-586, Pergamon Press.

Torgerson, W. S. (1961). Distances and ratios in psycho-physical scaling. *Acta Psychological* 19, pp. 201-205, Elsevier.

Veit, C. T. (1978). Ratio and subtractive processes in psycho-physical judgment. *Journal of Experimental Psychology: General* 107, pp. 81-107, American Psycological Association.

Zahedi, F. (1986). The analytic hierarchy process – a survey of the method and its applications. *Interfaces* 16, pp. 96-108, Informs.

Zahir, S. (2006). Eliciting ratio preferences for the analytic hierarchy process with visual interfaces: A new mode of preference measurement. *International Journal of Information Technology & Decision Making* 5, pp. 245-261, World Scientific.

DTU Transport performs research and provides education on traffic and transport planning. It advises the Danish Ministry of Transport on infrastructure, economic appraisals, transport policy and road safety and collects data on the transport habits of the population. DTU Transport collaborates with companies on such topics as logistics, public transport and intelligent transport systems.

DTU Transport
Department of Transport
Technical University of Denmark

Bygningstorvet 116B
DK-2800 Kgs. Lyngby
Tel. +45 45 25 65 00
Fax +45 45 93 65 33

www.transport.dtu.dk